

11 March 1982

Prospect of sanity for oil prices

The weakening of oil prices in the past six months may — just may — presage the ending of the long recession. Will the oil consumers have the wit to see what they must do now?

Is the recent decrease of the international price of oil a sign that the great economic recession is about to end, or the first swallow whose appearance does not mean that summer has begun? Of necessity, there can be no quick answer. Only time can tell. But there are serious dangers in the arguments now going the rounds among economic soothsayers in the industrialized oil-consuming countries that rising prosperity is once more just around the corner. The most obvious risk is that people will now be tempted to falter in their pursuit of the strategies that have led to the steady reduction of the consumption of petroleum from the members of the Organization of Oil-Exporting Countries (OPEC), and thus to the decrease of prices in the past few months. The second risk, less obvious, is that the oil-consuming states will continue to under-estimate the need — and the opportunity — for making arrangements among themselves for buffering the cost of oil against the fluctuations decreed by OPEC and others.

That the weakening of oil prices could be a cheerful sign is beyond dispute. Just as successive increases of the international price of oil since 1973 helped to create the blend of price inflation and economic deflation that has hamstrung the oil-importing states, so a reduction of price should have the opposite effect — less rapid inflation, more economic activity and people spending less on petroleum products and thus having more to spend on other things. The calculation is that a 10 per cent fall of oil prices should reduce the costs of goods and services by about one per cent and yet lead to an increase of demand in the oil-consuming states of about 0.5 per cent. The decrease in the price of oil in the past six months is more like 12.5 per cent than 10 per cent, but the economists are right to say that even the small reductions of prices and small increases of production that would result would feed on each other. In any case, the optimists argue, there may be further reductions of oil prices in the months ahead and so, with luck, a recovery of many industrialized economies.

Before throwing their hats in the air, the oil consumers should think harder than is their habit about what lies ahead. Although the decision by Saudi Arabia to reduce its oil production by 1 million barrels a day (50 million tonnes a year) will not directly affect the now weak international market for oil, it may ensure that the emergency meeting of OPEC provisionally arranged for 18 March is less incoherent than its immediate predecessors. For OPEC is far from dead, the jubilant obituaries of the past few weeks notwithstanding. But OPEC, commonly called a cartel, has yet to demonstrate that it can function at a time of falling demand for oil as a means of restricting output on a scale sufficient to influence the international price of oil. Its essential weakness is that many of its members have come to look to the whole of their oil revenues for their own economic support, with the result that they cannot with equanimity contemplate either a reduction of the price of oil or a reduction of their own exports. Apart from Saudi restraint, the only substantial restriction of exports that there has been so far has been caused by the determination of Iran and Iraq to do as much damage as they could to each other's oil installations as part of their private (and pointless) war. Yet even an agreement to a temporary reduction of exports, say for six months, could, if it were substantial, steady the falling price of oil and perhaps even make it rise again. The swallow that the industrialized oil importers have joyously been contemplating could quickly vanish.

More lasting economic forces within the oil-consuming states will also help to moderate the rosy prospect ahead. If the price of oil decreases, some plans for exploiting marginally economic reserves of oil and natural gas in the oil-consuming states will have to be postponed, moderating the rising trend of oil consumers' domestic production in the past few years. In any case, if industrial production does pick up again, so will the demand for oil. And if OPEC can find a basis for functioning as a true cartel, fixing the price and agreeing within itself not to break the rules even if exports should fall drastically as a result, its members' reduced demand for goods and services will partially offset the expected increase of economic activity among the oil consumers. So the economic recovery now possible is by no means a foregone conclusion. And to the extent that the oil consumers are by no means in charge of events, they will simply have to hope for the best in the next few weeks.

At the same time, they should recognize that whatever happens within OPEC will not bring back the circumstances of a decade ago, when the international price of oil was roughly a tenth of what it has become. Not merely has much of the oil that used to be extracted from the Middle East at a cost of \$1 a barrel or less already been sold, but the arrangements that have been made in the oil-consuming states to increase domestic production (Britain, in the North Sea, for example) provide a floor below which the price of OPEC oil need not fall. OPEC members would more readily settle on a rational policy in the next few weeks if they were given some kind of assurance that, even in present circumstances, none of their customers is looking for a return to the prices of the 1950s. Ever since 1973, the oil consumers have set their faces against anything that might smack of joint negotiation on the international price of oil — and they have been justified by the events of the past few months. Now, though, when the boot is on the other foot, it might make sense that the oil consumers should break silence to give their suppliers an intelligent account of what the next decade holds.

More of the same will be the best recipe. Only now, a decade after the price of petroleum increased fivefold in as many months, have the benefits of conservation become apparent in the energy budgets of the oil consumers. Both in the United States and, to a lesser extent, in Western Europe, industrial users of energy have understandably led the way. They have been encouraged to invest in energy-saving equipment by the high costs of oil and of equivalent fuels. Japan, almost wholly dependent on imported fuel, has done even better — again for reasons that are entirely understandable. But what the experience of the past few years has shown is that energy conservation is necessarily a slow process, one in which the cost of energy is only one of several influences. The doubt now raised by the falling price of oil is that the continued hankering of some governments, that of the United States especially, to shield people from the true cost of energy for political reasons will take the edge off the incentive to energy conservation. (Exhortation will not do instead, for although a barrel of oil saved is exactly equivalent to a barrel of oil extracted from the ground, saving oil when it would be cheaper to burn it is an economic waste.) The trend, apparent in the past few years, is to replace oil by other sources of primary energy. The modest reduction of the price of oil should not, however, take the wind out of the sails of nuclear energy programmes, or even schemes



for turning coal into alcohol or sunlight into electricity, provided they are strictly economic.

There are also more subtle steps to take. The oil consumers should have been heartened by the demonstration provided in the past few years that even when a commodity is inherently as scarce as crude petroleum, it cannot be indefinitely insulated from market forces. If the real price is too great, people will use something else or will use less energy. And while prices can be kept artificially high for a short time, such a state of affairs cannot persist. What the oil consumers should ask themselves is whether this time of relative plenty in the supply of oil should not be used as an opportunity for making the international market in petroleum more effective. One snag is that the past decade has persuaded too many of them that their domestic supplies of oil, however small, should be reserved for domestic consumption. This is the spirit in which the United States Congress has in the past insisted that Alaskan oil should not be exported direct from Alaska, and in which the British Parliament has repeatedly asked that oil from the British sector of the North Sea should first be landed in the United Kingdom. The long-term interests of both countries would be better served, however, by making sure that this substantial quantity of oil could be used as a way of making a true market price for oil.

Free broadcasting

The British government has been converted to telecommunications. The other side of the coin is freedom.

Technological developments, frequently subversive of accepted economic truths, are occasionally a challenge to social norms. The technique of *in vitro* fertilization as a means of producing children may be such a case. The new techniques of telecommunications are a more immediate challenge, especially where governments have been unwilling to acknowledge which way the wind has been blowing. This challenge will soon crop up in Britain, where the government announced last week that it would indeed take up the licence granted to it (and all other sovereign states) by the conference of the International Telecommunications Union in 1979, and launch a direct-broadcasting satellite. By all accounts, this development is part of a recent and belated conversion of the British government to the belief that satellite communications have a brilliant future and, more surprisingly, to the view that it will be a social benefit if most British dwellings can be "cabled up" — linked to some broadband distribution system capable of handling more like forty than four distinct video signals, as at present.

The belief that the time has come when a choice of television signals may be a social benefit flies in the face of the entire history of British broadcasting. While the government's conversion (see page 105) seems to stem from a calculation that there is money to be made by manufacturing equipment that can also be sold elsewhere, what it has forgotten is that all its predecessors have fought hard for the notion of public service broadcasting, the Reithian concept that electromagnetic waves are inherently different from other kinds of communication in that they are best used for the transmission of seemingly messages, educative ideally, if entertaining then, in the last resort, improving. The subliminal knowledge that the British Broadcasting Corporation is the world's outstanding practitioner of this art may have helped, as much as the soundness of its plans, to persuade the British government to give the corporation the first refusal of the licence to use two television channels on the first British direct broadcasting satellite in 1985.

That is the thin end of the wedge. The way technology has developed, television signals have become a kind of token contraband — if they exist, and people hanker after them, people will find some way of seeing the pictures that they represent. At the worst, they may have to buy a video-recorder and the cassettes to go with it. Or a domestic dish with the appropriate re-coding equipment. Either way, the new technology limits the degree to

which the authorities can control what people see or hear by means of broadcast signals. But how can this knowledge be squared with what successive British governments have been saying within living memory — that there must be central control not merely of the signals but of their content? Charters approved by parliament at present restrict what can be broadcast by both the British Broadcasting Corporation and its commercial counterparts. Both kinds of organizations, for example, are enjoined to strike a political balance in the content of what they broadcast and not to offend public taste. From time to time, they get into trouble, and questions are asked in the House of Commons. But this happens only infrequently. For the most part, all concerned are united behind the principle that broadcasting is a community service which, until the past few weeks, has even lent support to the doctrine put forward by the Annan committee three years ago that it would be wrong to sanction pay-television lest that should allow some people to see television programmes denied to others, that it should permit the broadcasting of indecencies and (under Labour governments) that it should allow commercial companies as yet unlicensed to make a profit.

That mask must now crack, although it is far from clear whether the present British government has accurately appreciated the difficulties it will face in converting a voting public from the present condition in which television signals are needlessly scarce to one in which all those who wish to send out broadcast signals, and who can pay the cost, will be allowed to do so. There will be fierce opposition from those who protest that a free-for-all will spell the end of quality in British broadcasting. (Less Shakespeare, no "Upstairs, Downstairs" or "Brideshead Revisited".) Sir Harold Wilson, once prime minister, will be only one of those who argue that providing more opportunities for would-be broadcasters will pave the way for what he called "cultural pollution" two years ago. Perhaps the British government is calculating that, with all the delay that there has been, the pent-up sense of opportunity among would-be investors in new broadcasting systems is so great that they will put up with the most restrictive fine print on whatever licences they are given.

The truth, however, is that even now, the British government is moving too slowly. And the accident that the International Telecommunications Union has allowed each sovereign state in the Old World the right to a single geosynchronous satellite is technically a red herring, an artificial spur to decision. For if a commodious cable system is economic for a country such as Britain, a single satellite will not be able to fill it with signals, which in any case could be distributed more economically in other ways. Satellites do, however, have the administrative advantage that it is possible to think of licensing those who wish to use them, while preventing those who wish to broadcast only by cable from putting their signals out. The trouble with such an arrangement is that it would be hopelessly uneconomic, and illiberal as well. It would be better that the British government should face at the outset the fact the creation of a national cable television system financed with private funds entails the abandonment of the old paternalism of the regulation of what is broadcast, however well that may have been accomplished.

For practical purposes, this is what has already happened in the United States. As on many other issues, Britain is being dragged along reluctantly, and a decade or so behind. As on previous occasions, the change when it comes will not be as traumatic as is feared. Elsewhere, however, things are even worse. In France, for example, the Mitterrand government, still flushed with its election victory, undertook to dismantle the apparatus of day-to-day control under which French broadcasting labours resentfully. Last July, the government's appointed committee suggested how this should be done. Now, the government seems to be having second thoughts. Governments that would not dream of censoring written communications plainly, but wrongly, think that electromagnetic signals are fair game for government licensors and inspectors. They should understand that they can no more embrace this new technology while keeping the old regulatory procedures than they could require horseless carriages to be preceded on the highways by a man with a red flag.

Legal shadow over laser patent

First court verdict for Gould claims

Washington

The simmering debate over who should be credited with the invention of the laser entered a new phase last week when a federal judge in San Francisco, spurred on by a New York technology investment company, upheld a patent issued to physicist Dr Gordon Gould as taking precedence over the original patent issued to Dr Charles Townes and Dr Arthur Schawlow in 1960.

Dr Townes shared the Nobel prize, with two Soviet physicists, in 1964 for his work in quantum electronics and microwave spectroscopy leading to the development of the laser or Light Amplification by Stimulated Emission of Radiation. Dr Schawlow shared the Nobel prize for his work as a chemist in the same area last year. Both are generally credited with putting forward the first conceptual design for a laser in a scientific paper published in *Physical Review* in December 1958 (112, 1940; 1958).

Dr Gould, however, who worked a few doors away from Dr Townes's laboratory at Columbia University in New York in 1957, claims that a notebook kept at the time demonstrates that he was the one to coin the word laser. He also claims that his conception, as described by his investment company backer, was that of an amplifying device, and should thus entitle him to credit for the conception of the laser as a whole.

It could not be established on Monday whether Dr Gould's claim extends to the basic process by which photons are amplified within a laser or to some other feature of the device. The substance of the claim appears to be the notion of creating a population of excited atoms, capable of stimulated emission, by exposing a gas or other material to an intense flash of light.

Dr Gould's claims for patent rights extending to an estimated 25–35 per cent of all lasers on the market, have been vigorously pursued by a New York-based company Refac Technology Development. With another company, Patlex of Pennsylvania, Refac has bought 80 per cent of any royalties generated by Dr Gould's patent for a sum claimed to be over \$2 million, and says it is now defending his claims to the laser against those of the scientific and technical establishment.

Until last week, the courts had been relatively unsympathetic. Dr Gould has failed several times to have revoked the patent granted to Townes and Schawlow in 1960, when both were working at Bell Laboratories, on the grounds that he had

been the first to conceive a complete working laser.

In 1977, however, the US Patent Office awarded Dr Gould a patent on his claim to have designed the amplifying device referred to in the patent. Since then, Refac has sought to obtain royalties from all companies producing optically-pumped lasers, and says that on the basis of a 1979 patent, also granted to Gould, it will now seek royalties on the use of such lasers.

Last week's decision, the first time that the validity of the patent has been tested in a federal court, was based on a suit filed by Refac against the Palo Alto-based company General Photonics claiming infringement of Gould's patent rights and demanding royalties on all lasers produced by the company since 1977.

Refac claims that Dr Gould should have received the technical — if not the scientific — credit for the first laser. The arguments

convinced Judge Samuel Conti and, as a result of his ruling in favour of Refac, General Photonics has agreed to pay the company 8 per cent of all its future sales.

Dr Arthur Schawlow, now professor of physics at Stanford University, has said that the US Patent Office was wrong to issue Dr Gould with the 1977 patent, and that Dr Townes "may have told Gould what he was doing" during conversations at Columbia in 1958. Dr Schawlow also says that drafts of the subsequent *Physical Review* paper were already circulating in the laboratory in August 1958, the date on which Gould wrote down some of his ideas.

Mr E.M. Lang, president of Refac, argues conversely that Gould spoke to Townes about his ideas on how to produce an amplifier that would make the laser work, and that it was Gould's ideas which were later incorporated into early laser designs for which Townes and Schawlow

UN university goes on tour

Paris, February

Global modellers, the inheritors of those who gave us *The Limits to Growth* in 1972, gathered here from 22 to 25 February to advise the Tokyo-based United Nations University on its preoccupation with the problems of agriculture, energy and development. The symposium, organized by Professor Maurice Lévy of the Marie and Pierre Curie University (otherwise Paris VI), seems not finally to have persuaded the university to take global modelling to its bosom, but the rector, the Indonesian Soedjatmakto, and his four vice-rectors undertook to brood about the problem when they are back in Tokyo.

Part of the interest of the occasion was that it showed how stimulating modelling techniques of socio-economic problems have proved to be, if only as a spur to understanding how one variable is related to another. Thus a model of the rural economy of Bangladesh was widely acclaimed, but the work of the International Institute for Applied Systems Analysis, of the Organization for Economic Cooperation and Development and of Dr Sam Cole (University of Sussex) also received high marks.

Professor Donella Meadows, who has been working on a model of the rural economy of New England since her collaboration on the original Club of Rome report, argued passionately at the symposium for a kind of global network of global modellers, linked together by communications satellites and thus able to exchange information or computational codes easily. The United Nations University cautiously withheld its blessing.

One of the problems of such occasions is that the participants are in two camps —

those who would construct models that address some tangible aspect of a problem and those who hold that complexity is of the essence. Some of the most confusing (and energetic) contributions to the discussion came from those who argued that "techno-economic" models purporting to account for, say, the effects of fuel prices on food production, were certain to be inadequate, given their neglect of "socio-political" considerations.

The interest of the United Nations University in these studies stems from its wish partially to focus its interest on the encouragement of economic development on what the symposium called "the energy-agriculture nexus". Inevitably, in a gathering of systems analysts, some argued that "food" would be a better variable than "agriculture".

The United Nations University is not so much a university as the late U Thant's creation of a United Nations development agency. Its chief source of funds is a pledge of \$100 million from the government of Japan. The university has no students, while all its employees are on short-term contracts. According to Soedjatmakto, negotiations are now under way for a permanent building in Tokyo, although this is unlikely to be built before the end of his term of office in 1985.

Lacking a permanent establishment, the university works chiefly by forming links with academic groups elsewhere, in industrialized and developing countries alike. It was encouraged, at the end of its four-day stand in Paris last month, to be told by the newly created minister of external affairs at the French Department of Education that French academics will be asked to collaborate with the university. ●

had been awarded the patent.

The scientific record is sufficiently ambiguous to allow for conflicting interpretations of the facts. Lang argues that the details of a proposed amplifier included in the *Physical Review* paper and the subsequent patent were shown not to work, and that Gould's ideas should therefore take precedence. Schawlow's response is that the existence of a working amplifier was implied in the paper, and that even though the specific solution suggested did not succeed, other lines of approach were suggested which proved successful. He also maintains that ideas about possible amplifiers were part of the "state of the art" at the time, and hence not eligible for patent protection on behalf of any one individual.

Refac is using last week's decision to bolster its claims on behalf of Gould. Its share price rose 12 per cent in value after the verdict had been announced. However, others are unconvinced; Dr Schawlow says that the case was poorly defended by General Photonics, which has already admitted that it does not have the money to mount an appeal.

More telling is likely to be a separate suit filed by Refac against Control Laser of Florida, a leading manufacturer of optical lasers. This suit was filed within a few days of the patent being granted in 1977, and has already attracted wide interest from other manufacturers (who once intended to join the suit in opposition to Refac, but then decided to withdraw for fear of being challenged on anti-trust grounds).

When the Control Laser case comes to trial, the company stands to lose a considerable amount of money if the verdict goes against it. Mr Robert van Roijen, the company's president, said last week that the major point of dispute was whether Refac's 1977 patent covered merely the optically-pumped amplifier described in the patent application, or whether — as Refac claimed — the patent could be taken to cover the whole apparatus.

Mr van Roijen would be willing to pay royalties on the amplifier, but denies that a laser patent is involved because "it would cost only a few thousand dollars". His arguments are expected to be backed by Dr T.H. Maiman, a director of Control Laser, who was the first to publish details of a working model of the laser (*Nature* 187, 493; 1960).

Looming on the horizon, however, is another suit which Refac has filed against a separate company for infringement of the "use" patent; in this case, General Motors has joined the proceedings on the side of the defendant.

Refac continues to characterize such disputes as a David-and-Goliath conflict. The companies maintain that Refac is using Gould's research to support a position that has been consistently rejected by the courts, and that the San Francisco verdict was, in Mr van Roijen's words, a "travesty" that is unlikely to survive the next legal round.

David Dickson

Pest research centres

Foreign labs shut

New Delhi

Accusations that the big powers conduct espionage or militarily oriented research under the guise of science collaboration in developing countries have again surfaced in the wake of the recent expulsion of an American scientist, Dr David R. Nalin, from Pakistan. The expulsion followed allegations that the United States aided Pakistan Medical Research Centre (PMRC) in Lahore which he headed was engaged in research on the use of mosquitoes in germ warfare. Six years ago another US funded mosquito control project in India was closed down by the government following similar allegations.

Dr Nalin denies the charge. In an interview he said the allegation was part of a Soviet smear campaign against the United States in retaliation against American accusations that the Soviet Union had indulged in germ warfare using mycotoxins in Kampuchea. Dr Nalin claims that his centre was infiltrated by left-wingers who organized strikes and spread rumours of a connection between PMRC and the Central Intelligence Agency. Nalin said that one member of his staff had been shown to have Soviet connections. He said

that a Russian, Iona Andronov, who was found one day in the centre, turned out to be a reporter for the Soviet magazine *Literaturnya Gazeta* which "exposed" the centre in an article that was picked up by the world press.

There is some evidence that there were doubts in government circles in Pakistan over Nalin's centre even before the latest accusation of impropriety. Knowledgeable medical sources in India say that Pakistani scientists have been unhappy about PMRC for quite some time. It seems that Nalin's centre had been warned not to open a phial of Japanese encephalitis virus that had been brought for an experiment when it was well known that the disease does not occur in Pakistan.

Nalin admits that his centre had been engaged in work on Japanese encephalitis, but says that the work stopped some time ago. He denies that the unit ever handled genetically manipulated strains of *Aedes aegypti* mosquitoes as alleged by the Soviet press. Nalin said the centre's work was mainly on two species of *Anopheles* mosquitoes that carry malaria, a major problem in Pakistan. According to Nalin, PMRC had conducted pilot studies on control of the malarial mosquitoes by the release of sterile males and had developed an efficient way of sexing the mosquitoes to make the technique effective.

Nalin is associate professor of international health at the University of Maryland, which set up the medical centre in Lahore in 1961. Before becoming director of PMRC, Nalin worked on diarrhoeal diseases in Bangladesh at another United States funded unit run by Johns Hopkins University. That unit was expelled from India in 1975 following uproar in parliament about its activities in Calcutta. Among the reasons for its expulsion were its link with the US biological warfare laboratory in Fort Detrick and the US Navy and the fact that its American staff made frequent trips to India's border areas.

According to Nalin there is a similarity between the allegations of germ warfare that led to his expulsion from Pakistan and those raised in the Indian parliament in 1975. The Indian unit was said to be engaged in the release of genetically manipulated strains of *Aedes aegypti*, the vector of yellow fever which does not exist in India. The experiments to control a vector of a non-indigenous disease raised a furore and the parliamentary committee alleged that the US experiments were part of a programme to develop yellow fever as a germ warfare weapon. The Indian government closed down the New Delhi research unit despite protests from the World Health Organization (WHO) under whose aegis it was set up.

Nalin is not the first American scientist to have been expelled from the Indian subcontinent. Dr Carl Taylor, head of the Division of International Health at Johns Hopkins University, has been told by the

Expulsion denied

Washington

The Pakistani embassy in Washington denied last week that Dr Nalin had been expelled from the country because of the allegation over his involvement in bacteriological warfare research.

The Minister of Information at the embassy, Mr M.I. Butt, said that it had been decided not to renew Dr Nalin's two-year contract as director of the Medical Research Center in Lahore after it expired last August because of what he described as Dr Nalin's failure to stick to procedural requirements for administration and research, and tension with other members of the centre's staff which eventually led several of them to resign. However he added that Dr Nalin had been allowed to stay in Pakistan until the end of January in order to complete a report on his research.

Dr Nalin, speaking from the University of Maryland, said that the future of the research centre was now uncertain, since applications for renewed funding from the Agency for International Development and the National Institutes of Health had been disrupted by his departure. He also said that the head of the department, Dr R.H. Baker, was expected to take over the temporary running of the centre until its future had been decided.

David Dickson

Indian government not to come to India in any professional capacity — although he is still allowed into the country as a tourist. Taylor ran a nutrition project at Narangwal in the Punjab for some 15 years from 1960. This project, also funded by the United States, was closed down following allegations that it had been spying on the nearby major Indian airbase at Halwara. Taylor, who was recently passing through Delhi, denies any such charges.

After the closure of the WHO unit in New Delhi and the projects run by Johns Hopkins University, the Indian government has become more strict in the scrutiny of all projects involving foreign collaboration. A central agency that includes the defence adviser has been set up to give foreign-aided projects security clearance. For the past three years the United States has been trying to mount an earthquake project in Shillong and a sedimentology project in the Bay of Bengal. But both projects have been turned down — Shillong is in the politically troubled north-east region while India is sensitive about allowing geological studies in the Bay of Bengal with its huge oil potential.

K.S. Jayaraman

Space communications

UK jumps in

Britain is to have a fully-operational direct broadcasting satellite by 1986, the earliest date foreseen in a Home Office report on direct broadcasting by satellite published just under a year ago. The government's approval for the project last week reflects the urgency with which it now views the exploitation of space and information technologies. It hopes that an early start to the project will give British industry the chance to demonstrate its capabilities and to compete in the world telecommunications market.

The satellite's two direct broadcasting channels have been awarded to the British Broadcasting Corporation, largely because its plans for using them are most advanced. It was also announced last week that British Aerospace, GEC-Marconi and British Telecom are to form a joint venture called United Satellites to build the satellite system. The three companies are now consulting with N.M. Rothschild & Son, the merchant bank, over ways of raising approximately £150 million venture capital for the project.

The BBC will lease its two channels, while British Telecom may use another for telecommunications. The plan is to go straight into a fully-operational system involving two satellites in orbit (one a spare) and a third waiting on the ground. According to British Aerospace, the satellite will be based on the ECS, a telecommunications satellite built by the European Space Agency for which British Aerospace was the main contractor. Earlier suggestions that the satellite should be

based on L-sat, the second generation of European telecommunications satellites, were abandoned partly because L-sat will be unnecessarily sophisticated and too late.

The BBC will finance its operations mainly out of subscriptions for the programmes carried on one of the channels. One channel will carry repeats of programmes already shown on its two conventional TV channels and the other will be a pay-TV channel showing recently-released feature films and the pick of programmes from foreign television.

Viewers will be able to tune into the satellite by means of small domestic dishes. They may also be able to receive programmes relayed by cable from community dishes. Government plans for cable systems are to be announced shortly, although the Home Office refused this week to comment on a report in last week's *Economist* that the government is urgently considering plans to lay optical fibres between major cities and coaxial cable between households.

Nevertheless, last week's announcements are evidence of the British government's anxiety that Britain should catch up with other industrialized nations, especially France and Germany, which have been quicker to acknowledge the challenge of space and information technologies. Other moves to exploit some of the lessons Britain has already learnt in space technology, for example through its membership of the European Space Agency, may be expected within a year or so. Officials at the Department of Industry are working on a plan to increase efforts in maritime communications and on a national earth resources programme, part of which would be Britain's contribution to the European Space Agency's proposed Earth Resources Satellite.

Whether these proposals will involve Britain launching its own satellites and just how much public money will be available to fund them, however, remains to be seen.

Judy Redfearn

Information service

Online Britain

Pergamon-InfoLine, the online scientific and technical information service of the Pergamon/BPC Group, was launched this week by Kenneth Baker, Minister of State for Industry and Information Technology. Infoline will not only be available in the United States but in Britain will represent the first commercial carrier of scientific and technological databases. Although a fledgling from the home nest — its London-based computer — the company wants to break into the markets of Europe and North America and become a "world force in the growing online information industry". A new system, Video PatSearch, is to spearhead its marketing initiative.

For an annual subscription of about

£7,000 plus online connection charges (£35 per hour), Video PatSearch will provide an in-house patent search service. The system, which would sit compactly on a large desk, is the first of its kind to link an online terminal with new video disc technology to enable users to retrieve both drawings from the original patent specification and text. At present the system, which will operate in North America and Europe, is geared to the 750,000 US patents issued since 1971.

Hitherto, a subject search of US patents on file at the US Patent Trademark Office has been possible only by using the US Patent Classification System. A search on the computer system may take only 10 to 15 minutes on the average. Video disc technology opens the way to online access to other databases covering subjects where graphics display is necessary.

InfoLine was first set up in 1978 by a consortium including the Department of Industry, the Chemical Society of London, the Institute of Electrical Engineers and the British Library. A long period of development produced only a limited number of databases and in 1980 the partners decided they were unwilling to invest further in the project. The Pergamon Group stepped in and saved InfoLine from liquidation. Several million pounds went into the setting up of an operation based on the advanced VAX 11/780 computer system.

InfoLine complements BLAISE, the online medical retrieval service of the British Library. The databases offered by InfoLine include CA Search, the Chemical Service Databases; Compendex, the engineering database; PIRA Abstracts, covering the paper, publishing, printing and packaging industries; Management and Marketing Abstracts and GeoMechanics Abstracts. All of these databases have been available in the United Kingdom and Europe, via EURONET, for some time. They will now be offered to the US market. PatSearch, the database of US patents, on the US market for two years, will now be offered worldwide from the London computer.

At the opening both Kenneth Baker and Robert Maxwell, chairman of Pergamon-InfoLine, seemed optimistic that InfoLine would flourish where other government-backed schemes had failed. The minister referred to InfoLine as a "tiny bird that escaped the net" (of nationalization).

Jane Wynn

Belgian science policy

Harsh criticisms

Brussels

The parlous state of Belgium's science policy reflects the political and economic problems facing the country. The universities, which shoulder the bulk of state-financed research, are suffering particularly badly from budgetary cutbacks, and there are loud calls for the government

to make radical changes in research funding. At the same time, the unequal way in which government-funded research centres are concentrated in the more prosperous Flemish half of the country is the source of bitter complaints from the French-speaking Walloons.

The present escalation of the language war comes in response to the threat of major closures in the steel industry of the already economically depressed Wallonia, and has led Walloons to seek out other injustices to attribute to the present Flemish dominated coalition government.

Belgium has three national research centres, the Centre pour Energie Nucleaire at Mol in Flanders and the Institut de Radioéléments and the Institut National des Industries d'Extraction, both in Wallonia. Eighty-seven per cent of all the government subsidies in question are spent at Mol if the employment created by the nuclear research centre in related industries is taken into account. The last coalition government attempted to redress the balance, at least in terms of energy research, by favouring the coal industry. But this still left 70 per cent of government research funds being spent on nuclear energy compared with 6 per cent on extractive industries. Philippe Maystadt, Science Minister in both the present and previous governments, seems to have done little to soothe Walloon anger since increased expenditure on nuclear energy and solar energy mainly benefits Flanders.

The bad blood between Walloon and Flemish scientists is overshadowed by the crisis facing universities throughout Belgium. A law passed in 1971 lays down that universities are allocated a certain sum of money for each student and out of this the university has to pay for administration, teaching and research. Maystadt ruled out specialist centres for postgraduate research as an unjustifiably expensive way of funding research for a small country.

Research is thus left largely under the control of the university authorities, loosely supervised by a ministerial and an interministerial committee for science policy. In 1981, the universities spent BFr 25,000 million (£290 million) compared with BFr 15,000 (£170 million) million at other national or international research centres. This has led the universities into a trap with administrative and other running costs rising above the price indexation system and the allocations per student falling below the index increase. The universities have therefore started to cut back on research spending.

With the policy of budgetary restraint being followed by the present government, the problem is worsening. Professor André Jaumott, ex-chancellor of the Université Libre de Bruxelles, wants universities to separate their teaching and research functions, with no budget restrictions.

Maystadt has other plans. His ministry has asked Belgian industry and the state

New broom in Spain

Barcelona

Professor Federico Mayor Zaragoza, the new Spanish Minister of Education and Science, is assessing the draft of a law for scientific and technical research which should provide a framework for scientific research in Spain, something that has been almost untouched since General Franco's time.

Professor Mayor is well versed in the higher education and science politics of Spain, having occupied many key posts: rector of the University of Granada, acting president of CSIC (the Spanish science research council), president of the "Comisión Asesora" (the main fund for research grants) and "sub-secretario" (vice-minister) of education and science. He is now professor of biochemistry at the Autonomous University of Madrid and director of the Institute of Molecular Biology which is a unit of the "Centro de Biología Molecular Severo Ochoa". He was elected MP for Granada in 1979, but later resigned to become deputy director-general of UNESCO.

According to the draft law, the government intends to specify who will formulate Spanish science policy, who will control the different levels of organization and how research will be financed. It puts research under the control of a secretary of state who will report directly to the prime minister. An advisory committee, mainly composed of scientists, will propose general plans and priorities and submit an annual budget for research to parliament.

No change will be introduced in the structure of research bodies such as CSIC but, to circumvent the fact that these bodies have only permanent staff, a new institution will be created, which will be a public company that will employ research staff and act as an auxiliary to research centres. This will also facilitate the exchange of scientists between research institutes and universities. The draft will be offered for discussion to research institutions before going to parliament. However, new elections will take place before March 1983 and it is unlikely that such a complex law would pass through parliament before then.

Pedro Puigdoménech

departments what their research requirements are for the coming five years. On the basis of this information, research contracts are being handed out mainly to medium-sized companies to undertake this sort of applied research. For 1980-82, BFr 1,000 million (£11 million) has been allocated on projects such as optic fibre technology for the Belgian PTT. Solvay, the Belgian chemical giant, should also benefit from funding for industry-oriented biotechnology.

Jasper Becker

Martial law in Poland

Fresh appeal

A former president of the Polish Academy of Sciences, Dr Janusz Groszkowski, has sharply criticized a group of more than 140 intellectuals and academics who last month sent to the Sejm (parliament) an appeal for the ending of martial law.

The letter, which was also addressed to the United Nations General Assembly and the UN Commission on Human Rights, said that the imposition of martial law was contrary to the right of every nation to self-determination, and the freedom to determine its own political status and to ensure its economic, social and cultural development. This principle, it said, was the basis of the renewal movement which sprang up in Poland after August 1980. The letter particularly deplored "the attempts to divide the nation, setting workers against soldiers and the militia, the blockade of the means of communication in the whole country, the brutal breaking of workers' strikes by the militia and army, the internment of many thousands of people in prisons and camps. Cultural life, education and learning, it said, are being "paralysed", and the media rendered powerless.

Dr Groszkowski did not disagree with the content of the letter, nor with the demand that the authorities should "put an end to this confrontation with their own people". He felt, however, that the fact that such a letter was sent to the Sejm — a body which during the Gierek regime had totally lost the confidence of the Polish nation and indeed of the Communist party — did not give legitimacy to it. The Sejm, in its present form, is a creation of the Gierek period, now being exposed as an era of distortions and corruption. To approach an institution where people linked to the Gierek era are still active, was, he felt, "a serious political error".

Dr Groszkowski's letter reveals how little the reforms of the past 18 months have affected those in high places. In September 1980, a letter from Dr Groszkowski prompted the academy to call for major reforms in Polish political and academic circles from the Sejm downwards, to clean up political patronage and restore honesty and fair dealing to public life. Any appeal from Dr Groszkowski carries considerable weight among Polish academics — in 1976 he resigned as president of the academy when he was unable to gain legal redress for Mrs Aleksandra Hankus of the Krakow Technical University. Mrs Hankus had suffered 11 years of official harassment and almost a year in prison for libel after protesting, in 1964, that her research results had been stolen by a person enjoying political patronage who then went on to gain a doctorate on the strength of her work.

Vera Rich

Weapons in space

High frontiers

Washington

America should exploit its lead in technology over the Soviet Union to pursue an aggressive space-based military and industrial strategy that would provide a "technological end-run" around the Soviet military threat, according to a report published last week by the Heritage Foundation, a conservative Washington think-tank.

Describing a scheme for integrating various space-based technologies, from orbiting battleships capable of shooting down Soviet nuclear missiles to vast solar-power satellites, the report suggests that exploiting what it calls the "high frontier" of space could solve many of the nation's industrial and energy problems while providing effective military defence.

Both the political and the technical communities in Washington are sceptical of the proposal. Many question its technical feasibility; others claim that the projected cost of \$50,000 million over 10 years is unrealistically low.

However, the Heritage Foundation hopes that its proposals will capture the imagination of those in Congress and the Administration prepared to back an ambitious, space-based military strategy, and it suggests that in the long term it could prove much cheaper than developing nuclear weapons.

The general approach has some influential support. Last week, for example, Dr Richard DeLauer, Under-Secretary of Defense for Research and Engineering, told members of Congress that he believed the Soviet Union could be deploying laser weapons in space as early as 1983, and might have elaborate space battle-stations by the 1990s.

Such speculation was dismissed as "non-sense" by several members of the Defense Science Board, who claimed that an operational Soviet space-based laser was at least ten years away. However, Defense Secretary Casper Weinberger last Thursday quoted Dr DeLauer in support of the Administration's request for sharply increased defence expenditures to offset the Soviet lead in areas such as laser weaponry. (President Reagan is asking for \$433 million for research into laser weapons for the fiscal year 1983, compared with about \$340 million this year, and only \$165 million two years ago).

The Heritage report suggests a three-layered defence against a Soviet missile attack. The first would consist of 432 satellites orbiting the Earth, each carrying conventional heat-seeking rockets, and capable of detecting and destroying a Soviet missile soon after launch. This would be backed up by satellites with more advanced weaponry capable of intercepting re-entry vehicles in mid-course. Finally US missile silos would be protected

by non-nuclear projectiles capable of blowing up incoming warheads.

Heritage officials claim this scheme would both defend the United States against Soviet missiles with a 95 per cent chance of success, and could protect Western Europe against intermediate-range missiles. Their report encourages the development of space-based industrial processing and of the solar power satellite, enthusiastically endorsed by many leading aerospace companies but reported on sceptically last year by both the National Academy of Sciences and Congress's Office of Technology Assessment.

Assessment of the technical feasibility of the proposed programme varies widely. There is general support for some of the less radical technology that would be involved, such as the missiles used for silo protection, but greater doubt about the advanced systems needed to detect and destroy Soviet missiles.

Other uncertainties surround management, cost and political acceptability. On the first, the foundation suggests a dedicated national effort comparable with the Manhattan Project which, it claims, could have an operational system in place within five or six years. On cost, it estimates a total of \$50,000 million, of which \$35 million would be redirected from other areas of the defence and intelligence budgets, and the remainder from the National Aeronautics and Space Administration. General Daniel Graham, who was military adviser to Mr Reagan during the 1980 election campaign and is a former director of the Defense Intelligence Agency, headed the nine-member panel which put together the report. He admitted that this figure might be unrealistically low, but added that "even if we are 100 per cent under, we are offering a strategic bargain".

The whole proposal, however, raises arms control problems that may mean indefinite delay. For example, it would fuel Soviet criticism that the space shuttle is essentially a military programme. The satellites might also be accused of violating the UN Space Treaty, signed by both the United States and the Soviet Union in 1967, which prohibits the stationing of weapons of mass destruction in space.

Finally any hint of a shift in military and strategic planning away from the MAD doctrine would resurrect the heated technical debates of the late 1960s over the adequacy of defence measures, at that time over anti-ballistic missiles. Critics such as Dr Kosta Tsipis of the Massachusetts Institute of Technology and Dr Richard Garwin of IBM have been quick to point out that any space-based weapons system based on sophisticated communications is highly vulnerable to enemy attack.

General Graham said that a copy of the report had been passed to the Office of Science and Technology Policy, which is preparing a report on the future of the American space effort, and whose director, Dr Jay Keyworth, has been keen

to explore ways of exploiting the capability of the space shuttle. Dr Keyworth, however, is not expected to show much sympathy for the more radical technical proposals being suggested by the Heritage Foundation. He has already resisted congressional pressure to build and operate high-energy laser battle-stations for space defence against ballistic missiles. Orbiting battleships are unlikely to gain more approval — although they may get on to the list of advanced military technologies which is expected to appear on the Agenda of the new Science Council. **David Dickson**

Bangladesh conference

Fertile minds

Dacca

Agricultural and rural development were dominant themes at the conference of the Bangladesh Association for the Advancement of Science held in February at Joydebpur, twenty miles north of Dacca. The conference was opened by Justice Abdus Sattar, President of Bangladesh, who urged scientists to bend their energies towards self-sufficiency in food, population control and the development and exploitation of indigenous resources.

Professor A. K. Aminul Huq, this year's president of the association, pleaded in his address for the planned development of manpower to meet the needs of the domestic agricultural industry and of those "man-power importing countries" in which people from Bangladesh find jobs. He said more use should be made of agricultural wastes as a source of energy and pleaded for more government support for science and technology, given the enormity of the problems with which these professions can assist.

More immediate assistance with agricultural development is likely to flow from an agreement now reached between the government and the Saudi Fund for Development. The fund will provide US\$80 million — part grant, part low interest loan — towards the cost of the Chittagong fertilizer factory, which should be completed in 1985.

This project will cost \$467 million, and the plant will have a daily output of 1,000 tons of ammonia fertilizer and 1,700 tons of urea fertilizer. It has been financed by the Asian Development Bank, the Overseas Economic Corporation of Japan, the Abu Dhabi Fund for Arab Economic Development, the Canadian International Development Agency and the Islamic Development Bank.

When signing the most recent loan agreement, the managing director of the Saudi Fund for Development, Mr Mohammed Abdullah Al-Sugair, said that he would be recommending that the fund should be more sympathetic to the needs of Bangladesh. The fund is supporting six projects at a total cost of 727 million Saudi rials (£116 million).

M. Kabir

CORRESPONDENCE

Flower power

SIR — It is to be hoped that the outcome of Indian production of paper from water hyacinths, *Eichornia crassipes* (Mert.) Solms (*E. speciosa* Kunth.), reported by Jayaraman¹, will be more successful economically than previous attempts to use the plant's fibre in this way.

His statement that "all efforts so far have been directed at eradication of the weed" overlooks some interesting developments elsewhere. It is only in slow-moving waterways that the species becomes a pest. Soon after its introduction into Hong Kong and South China, in the first decade of this century², came cultivation, primarily as a feed for pigs, taking full advantage of the rapid multiplication of the plant. It is grown in South China today, in ponds and flooded fields, as a feed for livestock and poultry³. Burnt, it can be a useful manure, because of its high potash content.

Where it is plentiful, and not required for other purposes, straw is probably preferable as a source of paper, because of its higher yield⁴.

WILLIAM GARDENER

Colchester, Essex, UK

1. Jayaraman, K.S. *Nature* **291**, 183 (1981).
2. Dunn, S.T. & Tutchter, W.J. *Flora of Kwangtung and Hongkong (China)* 6, 7, 281 (Bulletin of Miscellaneous Information, Kew; 1912).
3. *Iconographia Cormophytorum Sinicorum* Vol. 5 (Academia Sinica, Peking, 1976).
4. *Bull. Imp. Inst.* **24**, 267 (1926).

Wages of war

SIR — It would be a disservice to the scientific community if the assertions of R.G.S. Bidwell (*Nature* 11 February, p.452) were allowed to stand unrepudiated. This author clearly sees a direct link between the imposition of sanctions and the waging of war, and appears to welcome the opportunity of contributing to it. He refers disparagingly to the activities of pacifists in the Great Wars, not realizing that their efforts in the First World Disarmament Conference were within a hair's breadth of creating conditions which would make war an impossibility. Sanctions, on the other hand, can do nothing but increase the likelihood of war, because they only serve to nourish suspicions of malevolent intent on both sides. Never have Soviet scientists that I have spoken to expressed any desire to destroy us, as Bidwell is close to suggesting.

No scientist working for the benefit of humanity can surely wish to be associated with the notion that science is "ours", not "theirs", because we have paid for it. If the fruits of science are not made available to others, then it is of no earthly use. If the science we do in our laboratories is considered unsuitable for widespread application, it is better left undone. What we pay for is the prestige of making a discovery, the benefit to our industry of detailed technical knowledge and experience, and the chance to contribute to mutual understanding between nations, but not for the right to deny this knowledge to the rest of the scientific community.

Lack of understanding through insufficient information gives rise to some of the grossest inequities in the world today. The eclipse by the Polish situation of the tragic events taking place in El Salvador is a direct result of restriction of flow of information. It is

important that scientists should avoid involvement in such a practice, and refrain from endorsing the dubious and hypocritical moral judgements of our cold-war strategists.

ROGER R.C. NEW

Department of Biochemistry,
University of Liverpool, UK

Myth or fact?

SIR — Is the creation versus evolution argument really irreconcilable?

To those who cling to every word of the Bible nothing can be said. But surely the Old Testament and other ancient texts are the repository of ancient knowledge. Perhaps the creation is a distant "folk memory" — reinterpreted by succeeding generations — of events that actually took place at a time (15,000 yr BP?) beyond documented recall.

The geological evidence is that marked changes in the environment took place at the Pleistocene/Holocene boundary. Could the ancient accounts of the "beginning of the world" be telling us that the changes at this time were more profound than the geologists normally consider? Looking at ancient legends about the beginning of the world it is striking how the same motifs occur time and again: darkness over the face of the oceans; little or no dry land; the low heavens (the firmament) yet to be separated from the earth; a creator or creators usually of remarkably human kind.

Mythologies come down to us in garbled form; and modern science knows less than it cares to admit. So the split on this issue is difficult to understand.

There is a geological theory with a great deal of evidence in its favour that, in outline at least, can explain the drastic world changes alluded to in ancient legend. But for some reason it has little or no currency: for this reason one hesitates to elaborate . . .

R.B. TOPHAM

Abingdon, Oxon, UK

SIR — It seems to me that creationist science is at least as good as the evolutionary theology propagated by Jon Marks (*Nature* 28 January, p.276). To take his points in order: the Hebrew used in *Leviticus* 11: 19 and translated "bird" clearly denotes all warm blooded flying animals since the distinction between birds and mammals was unknown when it was written; Jesus did not ascend to paradise (or heaven) 3 days later as Marks claims, He ascended to heaven bodily at Pentecost 40 days later; what Jesus meant in Luke 23:43 was that both the thief and He would that day be in heaven in their spirits or souls. What Origen and Maimonides thought about the Bible is not pertinent; they were human and could err. There is no doubt, however, that Jesus Christ believed in a 6-day creation, a literal Adam, a real Garden of Eden, an actual world-wide flood and a real whale that swallowed and regurgitated Jonah.

As for the morality of the Bible, neither the rape of Dinah nor the seduction of Lot by his daughters is held up as a good thing. Rather they are consequences of the fall of Man, just as Dr Marks' cynical unbelief is. The execution of Sisera was surely justified for war crimes; he had oppressed Israel harshly for 20 years (*Judges* 4:3). We do not know all the details but no doubt there was something particularly appropriate that he should be killed by a woman with a tent peg. If Dr

Marks read his Bible with understanding he would realize that *Ecclesiastes* 3:9 is a perfect description of the futility of man without God, and one that fits Dr Marks himself.

It may be that someone who knows nothing but the 750,000 species of insects is a colossal bore, but someone who knows that, understands every language known to man, comprehends the whole of mathematics, physics and even anthropology would probably consider Dr Marks and his petulant fist waving at his Creator a little boring.

That chimpanzees and humans should be so similar genetically merely shows their Designer recognized a good thing when He saw it.

For Dr Marks to accuse creationists of fraud is a little thick. For whose benefit was the Piltown fraud perpetrated? What about Haeckel's faked photographs showing that embryology recapitulates phylogeny?

As for obscurantism; the word means the denying of inquiry. Is it not the evolutionists who hold that the theory of evolution is unchallengeable and must be accepted as fact?

T.J. HAMBLIN

Royal Victoria Hospital,
Bournemouth, UK

Culture conscious

SIR — In a recent issue (*Nature* **294**, 42; 1981) Adrienne Zihlman reviewed Sarah Hrdy's *The Woman That Never Evolved* and my own *The Evolution of Love*.

Professor Zihlman concluded that "Both of these books join the growing stacks of sociobiological attempts to integrate genes and human behaviour. They fail by ignoring the intervening levels. . . ." and that "the book that has not yet been written is one that. . . integrates culture and biology". In fact my second chapter contains an elaboration of the interaction of cultural and biological evolution, and my view that cultural evolution has probably "initiated or promoted in humans a far greater number of new behavioural trends than biological evolution has". My index shows that cultural evolution is discussed or mentioned significantly on at least 46 pages of the total of 291 pages of text.

SYDNEY L. W. MELLEEN

Vevey, Switzerland

SIR — In her review of my book *The Woman That Never Evolved*, Adrienne Zihlman quotes me out of context and implies that I state women's sexual activity is "assertive and temporarily insatiable". But I was referring to a monkey in oestrus, and was contrasting such monkeys with human females. I wrote: ". . . what earthly relevance does the conduct of this monkey have for understanding her culture-bearing cousin, whose solicitations are sedate, self-conscious, often elaborate in their subtlety and indirection?" (p.160).

My point was to show that cultural practices such as *purdah* and clitoridectomy, institutions such as marriage, and perhaps especially the myths and values that are a very real component of each human individual, have profound effects on the sexual behaviour of women (pp.179–187).

SARAH BLAFFER HRDY

Department of Anthropology,
Rice University,
Houston, Texas, USA

NEWS AND VIEWS

Creation, evolution and high school texts

from Robert M. May

ANYONE who has seen Clarence Darrow — as played by Spencer Tracy — triumph over William Jennings Bryan in the powerful movie *Inherit The Wind* will be pardoned for thinking that the Scopes Trial helped establish evolution over creation in the United States in the 1920s. This impression is, indeed, fairly widely held, and is explicitly affirmed in at least one university-level biology text.

To the contrary, Grabiner and Miller¹ have shown that the pressures exerted by fundamentalists in the 1920s were successful in bringing about the systematic removal of references to evolution and Darwin in high school biology texts, particularly in the Southern States. Before the Scopes Trial in 1925, several of the widely used texts discussed evolution¹: Hunter's *A Civic Biology* (American Book, 1914), from which Scopes himself taught, had a three page section on evolution together with other associated material elsewhere in the book; Gruenberg's *Elementary Biology* (Ginn, 1924) was outspokenly evolutionist, with several chapters devoted to the subject and classification discussed from an evolutionary standpoint; Moon's *Biology for Beginners* (Holt, 1921) had several chapters on evolution and a picture of Darwin as the frontispiece. Other successful texts were more reticent: Smallwood, Reveley and Bailey's *New Biology* (Allyn and Bacon, 1924), possibly the most widely adopted text of the time, gave about two pages to evolution and omitted any discussion of the origins of man; Clement's *Living Things* (Iroquois, 1924) gave evolution only brief mention; and Peabody and Hunt's *Biology and Human Welfare* (Macmillan, 1924) explicitly excluded evolution. Grabiner and Miller give a simple rule for identifying texts published in the decade following 1925: "Merely look up the word 'evolution' in the index or the glossary; you almost certainly will not find it". Concerned about the effect the Scopes Trial might have on sales of his text, Hunter rechristened the 1926 edition *New Civic Biology* and removed most of the explicit discussion of evolutionary ideas; the word 'evolution' is no longer in the index. The 1926 edition of Moon's *Biology for Beginners* has the frontispiece portrait of Darwin replaced by a cartoon diagram of the digestive system, although the substantial treatment of evolution in the

text is retained and 'evolution' remains in the index (at last disappearing, however, in the 1933 edition). The treatment of evolution in Smallwood *et al.*'s *New General Biology* (1929) is even more perfunctory than in their earlier *New Biology*, and again the entry for 'evolution' is removed from the index. The most widely used text in the 1930s was Baker and Mills' *Dynamic Biology* (Rand McNally, 1933), which discusses evolutionary ideas in the final chapter; the book², however, manages to avoid ever using the term 'evolution' and concludes with the extraordinary statement that "Darwin's theory, like that of Lamarck, is no longer generally accepted".

As emphasized by Simpson³, it is not so much that the Scopes Trial precipitated these changes, but rather that both the trial and the changes were consequences of the upsurge of anti-intellectual conservatism and biblical literalism during the period in question. The essential point, however, is that evolutionary biologists were deluded in seeing the drama enacted in the Tennessee courthouse as a significant victory. The real battles were being fought over state and local decisions about which textbooks to adopt for high school biology courses, and here creationists advanced on a broad front throughout the 1920s and 1930s.

We would do well to keep these facts in mind today. Even as the community of professional biologists takes comfort from Judge Overton's incisive and unequivocal verdict against the state of Arkansas' law mandating equal time for 'creation science' in the biology classroom, there are signs that new editions of major high school biology texts are being eviscerated.

As stressed by Nelkin⁴, a sociologist whose *Science Textbook Controversies and the Politics of Equal Time* (MIT Press, 1977) is the canonical book on this general subject, it is "too early to get a detailed reading on actual changes; they are just being implemented and publishers won't talk". But there are some clear signs, summarized recently by Zuidema⁵. The index to the 1973 edition of *Biology: Living Systems* (Charles Merrill) gave 17 lines of page references under the heading

'evolution'; in the 1979 edition, 'evolution' is indexed in 3 lines. Recent editions of three Harcourt Brace series texts omit all mention of Charles Darwin, and one excludes the index entry 'evolution' (though the subject is covered in the text). This is sad, as Simpson³ has praised Harcourt Brace for "consistently and effectively oppos[ing] anti-evolutionist pressure" in the earlier period around 1930–1960. Another brief survey⁶ shows the 1977 edition of an Otto and Towle high school biology text has diminished the coverage of evolution by one-third compared with the 1973 edition.

Taking a different tack, some texts now include material on the Genesis account of creation, or on creation myths generally^{5,6}. Books in which special creation now appears include two editions of *Biology: An Inquiry Into the Nature of Life* published by Allyn and Bacon in 1974 and 1977, a 1974 Smallwood and Green text, and the 1980 Houghton Mifflin text *Biology: The Science of Life*. These books can be characterized as teaching about creationism, but not supporting it. More controversial is *Biology: A Search for Order in Complexity* (Zondervan), written by Moore who is a 'born again' professor of natural science at Michigan State University and a founder of the Creation Research Society. The book was chosen as one of seven officially approved biology texts by the Indiana State textbook commission in 1975, and in two of Indiana's districts was the only ninth-grade biology text available to students. Later, in 1977, the book was barred from use in public schools in Indiana as sectarian; it was relegated to library use as a reference work in Texas, but remains approved by state commissions in Alabama, Georgia, Oklahoma and Oregon⁵.

Underlying much of this is the system whereby 19 of the 50 United States have prescribed selection systems under which textbooks are approved for adoption in schools. According to rules which vary from state to state among these 'adoption states', one or several books may be approved for school use at a given level⁷. These texts may then be provided to schools, paid for with tax dollars; alternative books may be used but they must be bought by the individual schools or by individual students. These 19 'adoption states' are preponderantly southern, but importantly include

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, New Jersey 08544.

California (which accounts for about 10 per cent of the national textbook market) and Texas (which currently has an annual budget of \$45 million for high school textbooks)⁷. Given the highly competitive nature of the textbook market, the pressures that this system — with all its political and populist overtones — puts on publishers is obvious. Typical, and understandable, is the hawking of one publisher about his company's text⁸: "evolution runs like a thread throughout, but is mentioned specifically only in the last chapter".

Why are creationists not a force in Europe? The reasons are many and varied, but one simple set of numbers helps to characterize the difference. Writing of "revived dogma and new cult", Martin⁹ observes that in Northern Europe, around three to five per cent of those nominally Protestant are found in church on any given Sunday. In contrast, the latest Gallup polls show 51 per cent of all American teenagers in church on any given Sunday — and at least one-third of these are fundamentalists.

Grabiner and Miller's conclusions¹

about the Scopes Trial era have a message for today. "Readers may choose their own villain in the story we have told. Like us, some will find the greatest culpability in the scientific community itself, for the large-scale failure to pay attention to the teaching of science in the high schools. Others will blame the textbook authors and publishers for pursuing sales rather than quality. Some will attach blame to the politicians who exploited anti-evolution sentiment to get into, or remain, in office . . . But whatever the lesson one wishes to draw from the history of biology textbooks since the Scopes trial, we think the story itself is worth knowing. That the textbooks could have downgraded their treatment of evolution with almost nobody noticing is the greatest tragedy of all." □

1. Grabiner, J.V. & Miller, P.D. *Science* **185**, 832 (1974).
2. Baker, A.O. & Mills, L.H. *Dynamic Biology*, 681 (Rand McNally, New York, 1933).
3. Simpson, G.G. *Science* **187**, 389 (1975).
4. Nelkin, D. Personal communication (1982).
5. Zuidema, H.P. *Creation/Evolution* **2**, 18 (1981).
6. *Science* **81**, 58 (December 1981).
7. Dahlin, R. *Publishers Weekly* **220**, 28 (1981).
8. Henig, R.M. *Bioscience* **29**, 513 (1979).
9. Martin, D. *Daedalus* **111**, 53 (1982).

wavelength respectively measured. The realization of the second through the Cs clock has a reproducibility of better than 1 part in 10^{13} . Despite the fact that wavelength comparisons can be made at an accuracy of parts in 10^{11} , the realization of the metre through the Kr lamp, which has a relatively broad spectral profile, is reproducible only to about 4 parts in 10^9 . As saturated absorption-stabilized lasers operating in the visible are internationally reproducible to a few parts in 10^{11} , there is obviously scope for an improved definition of the metre.

From what has been said, it might seem that a good redefinition of the metre would be to specify it, as at present, in terms of a wavelength, but to choose as this wavelength the most accurate that can be generated with a visible laser source. (For practical reasons it is easiest to use visible radiation for length measurement.) This would indeed be possible, but it would not be easy to decide which laser would be best for the definition. It would also mean that if a better, more accurate laser became available in a few years, then the definition might need to be changed again. It has therefore been decided to use the high precision of the frequency standard and to define the metre in terms of the second and a fixed value for the speed of light. The redefinition most likely to be recommended is that "the metre is the length equal to the distance travelled in a time interval of $1/299,792,458$ of a second by plane electromagnetic waves in a vacuum" (Comité Consultatif pour la Définition du Metre *BIPM*, 6th session, Sèvres; 1979). In this definition, c becomes simply a conversion factor between time and length.

This is a concept that we are familiar with through the use of the 'light year' as a unit of length. Both the proposed definition and the light year assume that the speed of light is constant but, as has been pointed out earlier (W. R. C. Rowley, *Phys. Bull.* **31**, 334; 1980), the constancy of c is also assumed implicitly in the present definition of the metre. What is fixed in the krypton atom is its energy level structure, and so the transition energy E . The wavelength emitted from the transition is given by $\lambda = hc/E$, where h is Planck's constant. The value of c to be used in the new definition has been chosen to be that measured with the existing metre, so that continuity is maintained.

The proposed redefinition of the metre will, however, rely heavily on laser sources for its realization. Any radiation of known frequency will automatically be of known wavelength, because c is fixed, and so may be used as a wavelength and length standard. And as the reproducibility of the frequency standard is improved, so the accuracy of the metre will follow it — which is good news for everybody, except for those who have always wanted to measure the speed of light but haven't got round to it yet. □

VELOCITY OF LIGHT TO BE CONSTANT — OFFICIAL

from Bridget Marx

ANYONE wanting to measure the speed of light had better do it soon. Indeed, it may already be too late. From the autumn of 1983 no one will be able to follow the example set by such illustrious people as Fizeau, Foucault, Maxwell and Michelson. Why is this? If the proposed new definition of the metre is put forward to and accepted by the General Conference on Weights and Measures, then it will entail the introduction of a fixed numerical value, $299,722,458 \text{ m s}^{-1}$, for the speed of light. Then 'c' will no longer be a measurable quantity but will be a fixed number.

The present situation is as follows. We have the second, the unit of time, defined in terms of a transition frequency near 9 GHz from a caesium atom: the second is the duration of 9,192,631,700 periods of this radiation. We have the metre, the unit of length, defined in terms of a transition wavelength at 605.8 nm from a krypton atom: the metre is 1,650,763.73 times the vacuum wavelength of this radiation.

Using these two defined units, we can measure the speed of light, in metres per second, in one of several ways. The earlier investigators used the technique of actually timing a pulse of light travelling over a measured distance. But light travels so fast that this is a difficult technique to use at high precision. More recently, accurate measurements of c have been made by

using the relationship between frequency, ν , and wavelength, λ , which holds true for all electromagnetic radiation: $\nu = c/\lambda$.

An accurate measurement of ν , in inverse seconds, and λ , in metres, results in an accurate experimental value of c . One difficulty is that the second and the metre are defined in terms of radiations at very different parts of the electromagnetic spectrum. These measurements therefore involve 'frequency-chains' from the 9 GHz Cs frequency to the 4.95×10^5 GHz Kr frequency — a series of oscillators the outputs of which form bridges to span the spectrum, so that frequencies in the optical region can be compared with those in the radiofrequency region. So for one particular radiation source, usually a laser, the frequency and wavelength can be found by comparison with those quantities which define the units of time and length, giving a precise experimental value of c .

What limits the precision of such an experiment? It is partly determined by how well you can do it, but also by the limiting accuracy with which the basic units can be realized. That is, it depends on how well Cs clocks and Kr lamps can be reproduced internationally, and their frequency and

Bridget Marx is at Coherent (UK) Ltd, Cambridge Science Park, Milton Road, Cambridge CB4 4BH.

Gravity waves seeding ionospheric irregularities

from Jürgen Röttger

THE upper atmosphere of the Earth, particularly above 50 km altitude, is partly ionized by solar radiation. The electron densities that result display diurnal, seasonal and other quasi-regular variations, the peak densities occurring within the so-called F-region of the ionosphere (see Fig. 1). Superimposed on these comparatively predictable features of ionospheric behaviour are disturbances such as those caused by solar flares and magnetospheric storms. Other irregularities, detectable by means of radio waves reflected or scattered from the ionosphere and unrelated to any obvious extra-terrestrial disturbance, are often difficult to explain. However, recent work has demonstrated that the neutral atmosphere plays a crucial role in their generation.

Their usual name apart, irregularities in the ionospheric plasma are rather common and regular features. Particularly near the equator, F-region irregularities occur quite frequently during the night and often exhibit a rather ordered, even periodical occurrence^{1,2}. Various plasma instabilities are responsible for generating irregularities with dimensions less than a few kilometres³ while dynamic processes in the neutral atmosphere are responsible for larger-scale periodical features of equatorial F-region irregularities¹⁻⁵. These processes are internal gravity waves which cause travelling ionospheric disturbances (TIDs) in the ionized part of the upper atmosphere. Recent research has shown that the disturbances can be amplified in the equatorial F-region due to mutual coupling with dynamic processes in the ionospheric plasma⁶. The resulting phenomena are periodic modulations, wave breaking as well as seeding or triggering plume-like irregularities.

Understanding of the equatorial F-region plasma irregularities has advanced in the present decade by the combination of theory and numerical simulations with data from ionosondes, backscatter radars, rockets and satellites and from analyses of scintillations of satellite signals¹⁻³. The irregularities, which are also called 'equatorial spread-F', commence after the bottomside profile of the post-sunset F-layer has steepened. Steep vertical gradients of plasma density can be unstable to the classic Rayleigh-Taylor instability because a heavy fluid — the plasma — is supported against gravity by the Earth's magnetic field. Perturbations in the plasma density of the bottomside F-region develop into density depletions and small-scale irregularities. Irregularities at the bottomside of the ionosphere occur in wave trains of patches

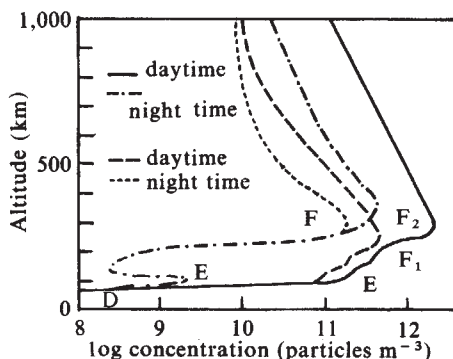


Fig. 1 Typical mid-latitude electron concentration profiles in the ionosphere. The larger daytime and night-time values occur at sunspot maximum, the smaller at sunspot minimum.

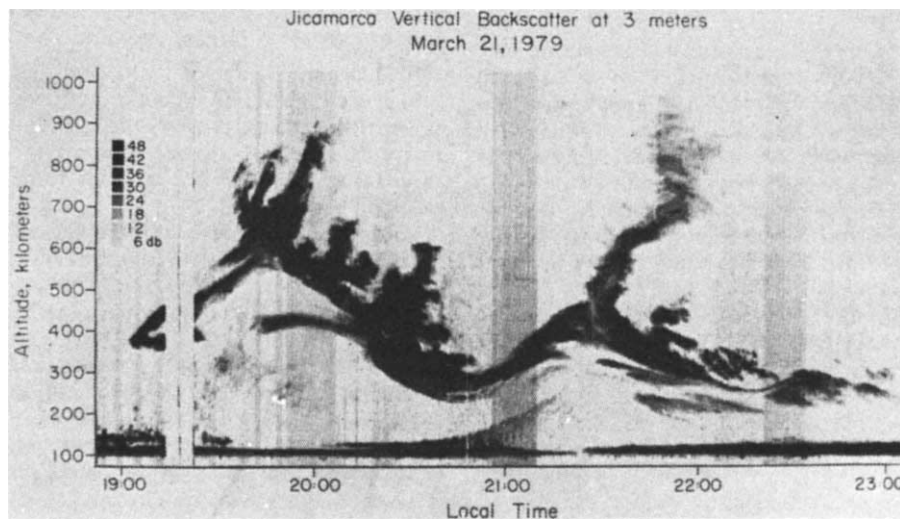
separated by some 100 km (ref. 4) and move eastwards and downwards. Plume-like irregularities are caused by the depletions rising, like bubbles in a fluid, to the topside of the F-region^{7,8}.

The new evidence⁶ that gravity waves initiate equatorial spread-F irregularities comes from observations carried out with the Jicamarca radar located close to the magnetic equator near Lima, Peru. The radar transmits at 50 MHz with 1 MW pulse power from a phased-array 300 m × 300 m antenna pointing nearly vertically. Backscattering from field-aligned spread-F irregularities of scale sizes of one-half the radar wavelength can be observed and radar pictures of the life history of spread-F irregularities recorded^{7,9,10}. One spectacular event recorded by Kelley *et al.*⁶ is shown in Fig. 2. The irregularities first occurred after sunset at the bottomside of the F-layer and, during the evening, the bottomside irregularities underwent nearly two full cycles of a quasi-periodic vertical motion. In both cycles, distinct plume-like structures extended to the topside of the F-layer and during downward motion, plumes were generated at intervals of about

20 min. These plumes are the signatures of plasma bubbles rising through the ionosphere up to altitudes of about 1,000 km. Because of their alignment to the Earth's magnetic field, the bubbles extend in stretched tubes to the north and south of the magnetic equator.

Since magnetospheric influences on the periodic bottomside modulation of the irregularity pattern can be ruled out, Kelley *et al.*⁶ argued that gravity waves have a role in this process in agreement with earlier investigations relating internal gravity waves to equatorial spread-F^{4,5,11,12}. This relation is rather intriguing since it allows a spatial resonance mechanism, as proposed by Whitehead¹³, to play a part — most effectively in the post-sunset equatorial F-region where the ionospheric plasma, drifting downwards at some 10 m s⁻¹, has similar downward phase velocities to those of gravity waves. An eastward plasma drift of 100–200 m s⁻¹ is also observed^{14,15}, matching typical horizontal phase velocities of gravity waves. When this matching occurs, the ionization pattern due to the wave stays in the same position relative to a drifting ionization parcel, and the wave-like perturbation in the ionization consequently increases. Even very small disturbances can be strongly amplified by this resonance and can reach the non-linear regime¹². Steep ionization gradients and depletions appear as a consequence. These were assumed to be the periodic source regions, that is, the primary perturbations necessary to initiate the Rayleigh-Taylor instability at the bottomside F-layer^{4,12}. As observed with radars, the generated ionization depletions or bubbles then drift upwards to the peak and the topside of the ionosphere^{7,8}. Independent of the radar observations, deep depletions (bite-outs with only some per cent density of the surrounding plasma) were detected by *in situ* measurements with rockets and

Fig. 2 Radar map showing echo traces due to backscatter from equatorial spread-F irregularities at altitudes larger than 200 km.



Jürgen Röttger is at the Max-Planck-Institut für Aeronomie, 3411 Katlenburg-Lindau, FRG.

satellites^{2,16}. The periodic features of spread-F plumes or bubbles also cause intense scintillation of satellite signals¹⁷ and adversely affect radio communication links.

Including the Doppler effect due to neutral background winds, Kelley *et al.*⁶ discussed the range of gravity waves which could match their data. They found that the long-period huge ionization perturbation observed in Fig. 2 would require a vertical velocity amplitude of 100 m s^{-1} — much too large to be realistic. They concluded that a gravity wave itself cannot explain the magnitude of disturbance but seeds the initial perturbations even if spatial resonance is only partially attained. Provided an initial finite amplitude perturbation exists due to gravity waves, they assumed that a nonlinear Rayleigh–Taylor growth for long-wavelength structures can be sufficiently high to explain the long-period modulation in Fig. 2. Smaller disturbances, with shorter periods of about 20 min, occurred during the downward motion and may be due to shorter-period gravity waves superimposed on the long-period perturbation.

These recent results establish that the life history of equatorial spread-F irregularities is a multistep process, starting with initial large-scale perturbations due to gravity waves which seed the generation by plasma instabilities of a cascade of smaller-scale

irregularities. As shown for the equatorial region, processes in the ionosphere cannot solely be treated from the viewpoint of plasma physics, but neutral atmosphere processes, namely internal gravity waves, have to be regarded as well. Since these waves presumably are generated in the troposphere^{6,18}, intriguing research on the coupling between the lower and upper atmosphere remains for the future. □

1. Fejer, B.G. & Kelley, M.C. *Rev. Geophys. Space Phys.* **18**, 401 (1980).
2. Kelley, M.C. in *Wave Instabilities in Space Plasmas* (eds Palmadesso, P.J. & Papadopoulos, K.) 291 (Reidel, Dordrecht, 1979).
3. Ossakow, S.L. *Rev. Geophys. Space Phys.* **17**, 521 (1979).
4. Röttger, J. *Zeitschr. Geophys.* **39**, 799 (1973); *J. atmos. terr. Phys.* **38**, 97 (1976); *J. atmos. terr. Phys.* **40**, 1103 (1978).
5. Booker, H.G. *J. atmos. terr. Phys.* **41**, 501 (1979).
6. Kelley, M.C., Larsen, M.F., La Hoz, C. & McClure, J.P. *J. geophys. Res.* **86**, 9087 (1981).
7. Woodman, R.F. & La Hoz, C. *J. geophys. Res.* **81**, 5447 (1976).
8. Ossakow, S.L., Zalesak, S.T., McDonald, B.E. & Chadurvedi, P.K. *J. geophys. Res.* **84**, 17 (1979).
9. Farley, D.T., Balsley, B.B., Woodman, R.F. & McClure, J.P. *J. geophys. Res.* **75**, 7199 (1970).
10. Tsunoda, R.T. *J. geophys. Res.* **86**, 139 (1981).
11. Beer, T. *Planet. Space Sci.* **21**, 297 (1973).
12. Klostermeyer, J. *J. geophys. Res.* **83**, 3753 (1978).
13. Whitehead, J.D. *J. geophys. Res.* **76**, 238 (1971).
14. Rishbeth, H. *Planet. Space Sci.* **19**, 357 (1971).
15. Woodman, R.F. *J. geophys. Res.* **75**, 6249 (1970); *Space Res.* **12**, 969 (1972).
16. McClure, J.P., Hanson, W.B. & Hoffman, J.H. *J. geophys. Res.* **82**, 2650 (1977).
17. Basu, S., McClure, J.P., Basu, S., Hanson, W.B. & Aarons, J. *J. geophys. Res.* **85**, 5119 (1980).
18. Röttger, J. *J. atmos. terr. Phys.* **39**, 987 (1977); **43**, 453 (1981).

Molecular basis of growth hormone deficiency

from Mike Wallis

THE first use of recombinant DNA technology to investigate the molecular basis of defective production of pituitary growth hormone (GH), one of the causes of dwarfism in man, is described in a recent paper by Phillips, Hjelle, Seeburg and Zachmann¹. Three related Swiss families were studied. The growth hormone deficiency is inherited² in an autosomal recessive mode; the six parents had normal stature but the growth of four of nine children was stunted. Using a cloned cDNA probe for human growth hormone (hGH), Phillips *et al.* showed that each of the four stunted individuals was homozygous for a deletion of at least 7.5 kilobases of DNA which includes most, or perhaps all, of the normal human growth hormone gene sequence. The dwarf individuals show a complete lack of GH (but not of other pituitary hormones) and when treated by the administration of hGH develop antibodies which neutralize its effect.

Phillips *et al.* took high molecular weight DNA prepared from blood samples from the dwarfs, other members of the same families and from control (unrelated) indi-

viduals and digested it with a variety of restriction endonucleases. The DNA digests were fractionated by electrophoresis, transferred to nitrocellulose filters (Southern blotting) and hybridized with a labelled hGH cDNA probe. For each restriction enzyme digest, several hybridizing bands were detected corresponding to various members of the family of genes known to be related to cDNA for hGH³. In each case, the affected (GH-deficient) individuals lacked one of the hybridizing bands, and their parents and most (but not all) of their siblings possessed less of this component (as would be expected for heterozygotes) than did normal individuals.

Interpretation of the data is complicated by the presence of at least two types of hGH gene (hGH-N and hGH-V; normal and variant) and probably at least three types of gene for the closely related hormone placental lactogen. The DNA sequence(s) coding for the placental lactogen are suffi-

ciently similar to those for hGH for extensive cross-hybridization to occur. Phillips *et al.* detected DNA fragments corresponding to at least the two hGH genes (hGH-N and hGH-V) and one placental lactogen gene, and their observations were apparently complicated further by the existence of an allelic polymorphism occurring in a placental lactogen gene. Their work suggests that the hGH-N and hGH-V genes are non-allelic, and that the hGH-V gene (which could produce a growth hormone differing from normal hGH at about 14 residues) is either non-functional or produces a poorly functional product in the GH-deficient individuals, all of which retain this gene. It is still possible, however, that the hGH-V gene is expressed and functional in normal individuals; deletion of DNA associated with the hGH-N gene in the dwarfs might lead to reduced expression of the hGH-V gene.

The complexity of the organization of the human GH and placental lactogen gene family contrasts with the situation in non-primate mammals. In the rat, for example, cDNA to rat GH cross-hybridizes with only a single growth hormone gene, the structure of which has been elucidated in detail^{4–6}. The gene contains four introns. The related protein hormone prolactin is also represented by a single gene in the rat genome^{4,7}; again, the gene contains four introns, the positions of which correspond to those in growth hormone, but the sizes of which are far greater (on the average by a factor of five). It seems likely that the genes of the hGH–human placental lactogen family underwent a period of rapid evolution, including gene duplication and multiple substitutions, during the evolution of the primates; not only does the organization of the genes for GH-like proteins in primates differ markedly from that in non-primates, but the amino acid sequences of hGH and non-primate GHs are also very different⁸.

The kind of approach used by Phillips *et al.*² should be applicable to the investigation of other types of hGH deficiency. Much human dwarfism, however, does not result from simple lack of GH. Other causes of short stature are recognized, but many children are very short despite normal immunoreactive GH levels and the absence of any other obvious cause. Recently, Rudman *et al.*^{9,10} have proposed that a substantial proportion of such 'normal-variant short stature' (NVSS) children are short because they have a defective GH, which is still detectable by radioimmunoassay. Growth may be inducible in such children by exogenous GH. This work has aroused great interest in the medical journals, though some aspects of it remain controversial^{11–13}. Application of recombinant DNA techniques should help elucidate the nature of such GH defects.

Dwarfism caused by lack of GH can often be treated by injection of hGH (but not non-primate GH) and such treatment

Mike Wallis is in the School of Biological Sciences at The University of Sussex, Falmer, Brighton, Sussex BN1 9QG.

has been applied effectively in the UK and many other countries using GH extracted from human pituitary glands. The supply has barely met the demand, however, and recently increasing demand and falling supplies have aggravated the situation (see *Nature News* 294, 200; 1981). Application of recombinant DNA technology has led to the production of hGH from bacteria^{14,15}, and clinical trials with such material are starting. If the prevalence of GH-treatable NVSS children is as high as Rudman and his colleagues suggest (perhaps two to ten times as many as those now treated with GH), the need for hGH produced by genetic engineering will be considerable. But the hGH produced in bacteria is unlikely to be completely without problems. The material currently being tested has a structure slightly different (due to an additional N-terminal methionine

residue) from the pituitary-derived hormone, and there is no guarantee that bacterially derived material will be substantially less expensive than that obtained from pituitaries. (Complete treatment of a GH-deficient child currently costs about £10,000.) □

1. Phillips *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 6372 (1981).
2. Rimoin *Birth Defects: Orig. Artic. Ser.* **12**, 15 (1976).
3. Fiddes *et al.* *Proc. natn. Acad. Sci. U.S.A.* **76**, 4294 (1979).
4. Chien & Thompson. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4583 (1980).
5. Barta *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 4867 (1981).
6. Page *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 4867 (1981).
7. Gubbins *et al.* *J. biol. Chem.* **255**, 8655 (1980).
8. Wallis *J. molec. Evol.* **17**, 10 (1981).
9. Rudman *et al.* *J. clin. Endocr. Metab.* **51**, 1378 (1980).
10. Rudman *et al.* *New Engl. J. Med.* **305**, 123 (1981).
11. *Lancet* **ii**, 399 (1981).
12. Crawford *New Engl. J. Med.* **305**, 163 (1981).
13. Prececc *Br. med. J.* **283**, 1145 (1981).
14. Goeddel *et al.* *Nature* **281**, 544 (1979).
15. Olson *et al.* *Nature* **293**, 408 (1981).

St Patrick and the bacteria

from J.S. Jones

THERE are two models of the absence of snakes from Ireland, neither of which, as the Creationists delight in reminding us, is directly testable. The older of these has it that St Patrick banished into the sea all manner of noxious and creeping things from his adopted land. More recent theories — such as MacArthur and Wilson's equilibrium theory of insular biogeography¹ — see this absence as a result of the unique faunal history of islands, which is profoundly influenced by their size and isolation.

The classic studies in island biogeography have been of birds but, surprisingly, more direct opportunities to test MacArthur and Wilson's theory are now coming from investigations of the patterns of bacterial and viral infections in man. Not only does it appear that the spread of disease in isolated islands of human population accords well with the theories of biogeographers, but even changes in the bacterial flora of a single individual and an 'archipelago' of his family, friends and pets follows the same ecological rules as those thought to control the biology of island birds.

The equilibrium theory of island biogeography

The MacArthur and Wilson theory attempts to predict the number of species on an island from a balance between immigration and extinction. Immigration depends largely on the island's distance from a mainland source. The rate of extinction is related to the island's area; the smaller the island, the fewer the members

of a particular species which it can support and the greater the chance that the species will become extinct through random fluctuations in population size. Considerable turnover will thus be expected; although the number of species on an island will be constant, their identity will change with time, with the accidents of immigration and extinction leading to a dynamic equilibrium.

In an important new book, Williamson² takes a critical look at some of the many attempts which have been made to interpret the biology of islands in terms of the equilibrium theory. He points out that although it has been known for more than a century that there is a relationship between the number of species and an island's area and isolation, the evidence for a constant random turnover on islands is much less convincing. A forty year study of the birds of Skokholm, for example, shows that there has been almost no species turnover during this period. Other claims³ that the population history of island birds in the Pacific and elsewhere supports the MacArthur-Wilson view of species turnover are based on very inadequate information⁴.

Can more adequate data be taken from the huge body of information on the ecology of infectious disease? If infection is regarded as the immigration of a disease and cure as its extinction then it seems that it may.

Humans on islands

Measles occurs only in man and, because immunity follows infection, the virus can persist only in a large population where there is a continuous supply of susceptible persons⁵. A study of 19 island populations⁶ shows that measles periodically dies out in

all those containing less than 500,000 people, and that (exactly as claimed by MacArthur and Wilson) the probability of extinction of the virus is greater on small islands than on large. The rate of immigration also has a major effect on persistence and turnover; in Iceland (where records of the incidence of measles date from 1869), there were intervals of up to seven years between epidemics before 1938, a period during which there was little movement of infected individuals into the island. After 1945, when immigration was common, the time between epidemics was never more than two years⁷.

A similar pattern of turn over has been traced in an isolated Eskimo village from the persistence of specific antibodies in those who survived polio infection⁸. One strain passed through the community 25 years before the investigation; another 15 years before, and a third only within the previous year.

Black⁹ points out that most primitive human communities exist as islands of population, and that many of their diseases show such patterns of turnover, with a balance between the immigration and extinction of pathogens which is related to each island's size and isolation. In Amazonian tribes, for example, the incidence of particular infectious diseases differs greatly (and apparently at random) between the various isolated communities. In different villages the proportion of the population with antibodies to measles varies from 0 to 98 per cent, to rubella from 0 to 95 per cent and to parainfluenza B from 0 to 89 per cent. Disease incidence again results from a dynamic equilibrium between a chance immigration and a period of persistence followed by extinction.

Humans as islands

To bacteria, all men are islands, safe refuges in an inhospitable sea. The human intestine is such a favourable habitat that every healthy individual has ten times as many symbiotic bacterial cells as he has cells of his own¹¹. Gel electrophoresis of a sample of enzyme loci makes it possible to detect individual strains of bacteria and provides a new and intriguing chance to study the island biogeography of man. Isolates of *Escherichia coli* from humans, water supplies and laboratory stocks are made up of a great number of genetically distinct and persistent clones, each of which has a unique enzyme profile¹². Gene exchange among clones is a very rare event¹³ so that *E. coli* is in some senses a complex of a very large number of genetically well differentiated 'species' that can be identified using biochemical markers. One clone, found in 1979 in a Massachusetts infant, is identical to the laboratory strain K12, first isolated 50 years ago.

Caugant *et al.*¹⁴ have studied the ecology of such clones (or species) within a Dr Bruce Levin of the University of Massachusetts. Over eleven months (which represents thousands of bacterial gener-

J.S. Jones is in the Department of Genetics and Biometry, University College London, Gower Street, London WC1E 6BT.

actions) this human 'island' supported a fauna of 53 clones. Most were transient, and appeared on only one sampling occasion but one persisted for several months and two others occurred for most of the period of observation.

The *E. coli* clones therefore showed considerable 'species' turnover although the actual number of clones remained fairly constant. New clones probably arrived by immigration from the food as the nearby islands in the Levin Archipelago (the two Doctors Levin, their children and their cat) shared a similar array of clones. Clones were lost by random extinction. These observations further support the view that the species composition of islands involves considerable flux.

Biogeography and disease

Other useful analogies may be drawn between the biology of island birds and that of human pathogens. For example, islands — such as the Galapagos — are uniquely favourable in promoting the origin of new bird species. It is possible that the evolution of new pathogens takes place primarily in highly sub-divided human populations and that the recent expansion of mankind to form a single large continental mass with

free interchange among its parts has greatly retarded the evolution of new disease organisms.

Birds on isolated islands with few competing species and slow species turnover have habitat preferences quite different from mainland populations². The ecology of pathogens in isolated human populations might also differ from that in modern cities.

After defaunation by volcanic explosions or poison gas, islands show changes in the diversity, turnover and ecology of their inhabitants^{17,18}; the return to equilibrium may be long delayed. Recolonization of a human host after antibiotic treatment could lead to similar and unexpected changes in the ecology of the pathogens which remain.

The most recent models of island biogeography use differences in the efficiency of migration and resource utilization of individual species to predict the species composition of birds on islands¹⁹. Clones of *E. coli* also differ greatly from each other in their ability to utilize resources²⁰. Developments in theories of island biogeography may therefore help to explain why some clones

are resident and some transient in the intestine¹⁴, and why certain pathogens (such as hepatitis B or herpes virus) are present in all human populations — including those on islands — while others occur only sporadically⁸. It is even possible that such developments may help to eliminate pathogens by means of the secular miracles of medicine rather than the direct intervention of a saint. □

1. MacArthur, R.H. & Eilson, E.O. *The Theory of Island Biogeography*. (Princeton University Press, 1967).
2. Williamson, M. *Island Populations* (Oxford University Press, 1981).
3. Diamond, J. *Science* **184**, 803 (1974).
4. Simberloff, D. & Connor, E.F. *Am. Nat.* **118**, 215 (1981).
5. Bartlett, M.S. *J. R. stat. Soc. A120*, 48 (1957).
6. Black, F.L. *J. theor. Biol.* **11**, 207 (1966).
7. Cliff, A.D. & Haggett, P. *J. Hyg., Camb.* **85**, 451 (1980).
8. Yorke, J.A. *et al. Am. J. Epidemiol.* **109**, 103 (1979).
9. Black, F.L. *Science* **187**, 515 (1975).
10. Hassell, M.P. *Oikos* **35**, 150 (1980).
11. Savage, D.C. *Rev. Microbiol.* **31**, 107 (1977).
12. Selander, R.K. & Levin, B.R. *Science* **210** 545 (1980).
13. Levin, B.R. *Genetics*, (in the press).
14. Caugant, D.A., Levin, B.R. & Selander, R.K. *Genetics* **98**, 467 (1981).
15. Levin, B.R. personal communication (1982).
16. Holmes, J.C. *Can. J. Zool.* **51**, 333 (1973).
17. Simberloff, D.S. *Science* **194**, 572 (1976).
18. Rey, J.R. *Ecol. Monogr.* **51**, 237 (1981).
19. Gilpin, M.E. & Diamond, J.M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 392 (1981).
20. Dykhuizen, D. & Davies, M. *Ecology* **61**, 1213 (1980).



100 YEARS AGO

ON THE CONSERVATION OF SOLAR ENERGY¹

The question of the maintenance of Solar Energy is one that has been looked upon with deep interest by astronomers and physicists from the time of La Place downwards.

The amount of heat radiated from the sun has been approximately computed by the aid of the pyrheliometer of Pouillet and by the actinometers of Herschel and others at 18,000,000 of heat units from every square foot of its surface per hour, or, put popularly, as equal to the heat that would be produced by the perfect combustion every thirty-six hours of a mass of coal of specific gravity = 1.5 as great as that of our earth.

If the sun were surrounded by a solid sphere of a radius equal to the mean distance of the sun from the earth (95,000,000 of miles), the whole of this prodigious amount of heat would be intercepted: but considering that the earth's apparent diameter as seen from the sun is only seventeen seconds, the earth can intercept only the 2,250-millionth part. Assuming that the other planetary bodies swell the amount of intercepted heat by ten times this amount, there remains the important fact that $\frac{2499999}{225000000}$ of the solar energy is radiated into

space, and apparently lost to the solar system, and only $\frac{1}{225000000}$ utilised.

Notwithstanding this enormous loss of heat, solar temperature has not diminished sensibly for centuries, if we neglect the periodic changes, apparently connected with the appearance of sun-spots that have been observed by Lockyer and others, and the question forces itself upon us how this great loss can be sustained without producing an observable diminution of solar temperature even within a human lifetime.

Amongst the ingenious hypotheses intended to account for a continuance of solar heat is that of shrinkage, or gradual reduction of the sun's volume suggested by Helmholtz. It may, however, be urged against this theory that the heat so produced would be liberated throughout its mass; and would have to be brought to the surface by conduction, aided perhaps by convection; but we know of no material of sufficient conductivity to transmit anything approaching the amount of heat lost by radiation.

Chemical action between the constituent parts of the sun has also been suggested; but here again we are met by the difficulty that the products of such combination would ere this have accumulated on the surface, and would have formed a barrier against further action.

These difficulties have led Sir Wm. Thomson, following up Mayer's speculation, to the suggestion that the cause of the maintenance of solar temperature might be found in the circumstance of meteorolites falling upon the sun from great distances in space, or with an acquired velocity due to such fall, and he shows that each pound of matter so imported would represent a large number of heat units depending upon the original distance. Yet the aggregate of material that would thus have to be incorporated with the

sun would tend to disturb the planetary equilibrium, and must ere this have shortened our year to an extent exceeding that resulting from astronomical records and observation. In fact, Sir William Thomson soon abandoned the meteoric hypothesis for that of simple transfer of heat from the interior of a liquid sun to the surface by means of convection currents, which latter hypothesis appears at the present time to be supported by Prof. Stokes and other leading physicists of the day.

But if either of these hypotheses could be proved we should only have the satisfaction of knowing that the solar waste of energy by dissipation into space was not dependent entirely upon loss of its sensible heat, but that its existence as a luminary would be prolonged by calling into requisition a limited, though may be large, store of energy in the form of separated matter. The true solution of the problem will be furnished by a theory, according to which the radiant energy which is now supposed to be dissipated into space and irrecoverably lost to our solar system, could be arrested and brought back in another form to the sun itself, there to continue the work of solar radiation.

CAN Mr Wallace throw any light on Mr Allen's somewhat extraordinary sentence: "I feel a genuine respect for every donkey I meet, when I remember that it was the mere accidental possession of an opposable thumb that gave my ancestors a start over his in the race for the inheritance of the earth towards the very close of the tertiary period." I take Mr Allen to be an evolutionist, but there is no place for accident in evolution, or in any other scientific theory. The "opposable thumb" must be the result of some conditioning factor, and this being so the word accident is quite out of place.

From *Nature* **25**, 440; March 9, 1882.

¹ Paper read at the Royal Society, March 2, by C. William Siemens, D.C.L., L.L.D., F.R.S., Mem. Inst. C.E.

REVIEW ARTICLE

Molecular mechanisms of variation in influenza viruses

R. G. Webster*, W. G. Laver†, G. M. Air† & G. C. Schild‡

* St Jude Children's Research Hospital, Memphis, Tennessee 38101, USA

† The John Curtin School of Medical Research, Australian National University, Canberra, ACT 2601, Australia

‡ National Institute for Biological Standards and Control, London NW3 6RB, UK

Influenza is unique among the viruses in its capacity to vary. Because of antigenic variation, it has proved impossible to control influenza by vaccination, and variation in virulence, host range and transmissibility influence the spread and severity of influenza epidemics. This variation is caused by sequence changes in the genes of the virus and this review summarizes recent information on the structure of the genes and their products, the way in which these vary and the effects of the changes on the biological activities of the virus.

OVER the past decade influenza epidemics have been relatively mild, causing only modest increases in morbidity and mortality. However, the potential of the influenza virus as a major and unpredictable threat to public health is highlighted by historic events. During the weeks from 19 October 1918 to the end of February 1919, an influenza A virus which, from circumstantial evidence, was almost certainly antigenically closely related to 'swine' influenza A virus, killed at least 20 million persons and affected perhaps one hundred times that number. It was without doubt one of the most devastating plagues ever documented. Such an event could happen again.

In 1979–80, ~20% of the harbour seal (*Phoca vitulina*) population of the north-east coast of the United States died of a severe respiratory infection with consolidation of the lungs, typical of primary viral pneumonia¹. Influenza virus particles were found in high concentrations in the lungs and brains of the dead seals. Antigenic analysis showed that this virus was closely related to fowl plague virus (A/FVP/Dutch/27 (H7N7)), a highly lethal influenza virus of chickens not previously found in mammals². What would happen if such an event occurred in man instead of seals? Would the resulting pandemic be similar to that of 1918–19?

Properties of influenza virus

Influenza viruses are classified into three types—A, B and C—on the basis of their type-specific nucleoprotein and matrix protein antigens. Type A influenza viruses are further classified into subtypes based on the antigenic characteristics of their surface antigens, haemagglutinin (HA) and neuraminidase (NA). Twelve distinct HA subtypes and nine NA subtypes are now recognized in the nomenclature system for influenza A viruses recommended by the World Health Organization³, and a thirteenth HA subtype has recently been identified⁴. Examples of the use of this nomenclature system are as follows: A/Hong Kong/68 (H3N2) is an influenza A virus isolated from man in 1968 with HA of subtype H3 and NA of subtype N2, while A/equine/Miami/63 (H3N8) is a virus isolated from horses, with HA of the same subtype as that of the human virus but with a distinct NA antigen. It is implicit in this system that all viruses with common H or N subtype have HA or NA antigens shown to be related by conventional laboratory tests (such as double immunodiffusion) while such relationships do not occur between subtypes. However, included among viruses of a common subtype designation will be strains showing considerable degrees of variation. Thus, the human virus A/Bangkok/1/79 (H3N2) contains HA and NA antigens demonstrably related to those of A/Hong Kong/68 virus but showing considerable degrees of difference from them. During

the present century three distinct subtypes of influenza A virus have caused pandemics in man—the H1N1 subtype which probably appeared in 1918, the H2N2 ('Asian') subtype which appeared in 1957 and the H3N2 ('Hong Kong') subtype which appeared in 1968.

This review will be concerned with type A influenza, which has been responsible for the major pandemics and has been investigated more extensively than types B and C.

The highly pleomorphic particles of influenza virus are enclosed by a lipid envelope, derived from the plasma membrane of the host cell, to which the HA and NA antigens are attached by short sequences of hydrophobic amino acids at one end of the molecules^{5–7}. Both surface antigens are glycosylated and some of the carbohydrate side chains possess antigenic activity characteristic of the host cell in which the virus grew.

Within the lipid envelope lies the matrix protein, which is believed to be structural in function. Within the matrix shell are eight single-stranded RNA molecules of negative sense (that is, the virion RNA is complementary to the messenger RNA) associated with a nucleocapsid protein and three large proteins, P1, P2 and P3, responsible for RNA replication and transcription. At least three virus-encoded non-structural proteins are found in infected cells, but their functions are unknown^{8–10}. Table 1 summarizes current information on the eight genes of the influenza virus and the molecular weights (M_r) of the proteins encoded by them.

Variation in haemagglutinin

The HA accounts for about 25% of the virion protein. It is responsible for attachment of the virus to cells and for the stimulation of antibodies that neutralize the virus. The HA monomer is synthesized as a single polypeptide chain which undergoes post-translational cleavage at three places. An N-terminal signal sequence is removed and, depending on the host cell and virus strain, the molecule is cleaved, with the removal of one or more intervening residues, to give two polypeptide chains called HA1 and HA2, with molecular weights of 36,000 and 27,000, respectively (Table 1). A sequence of hydrophobic amino acids near the C-terminus of HA2 serves to anchor the HA in the virus membrane.

The three-dimensional structure of Hong Kong (H3N2) HA has recently been determined¹¹ and it seems likely that HA molecules from other subtypes will have similar structures. The HA glycoprotein of influenza virus is a trimer built of two structurally distinct regions: a triple-stranded coiled-coil of α -helices extend 76 Å from the membrane and a globular region of antiparallel β -sheet that contains the receptor binding site; the variable antigenic determinants are located on top of this

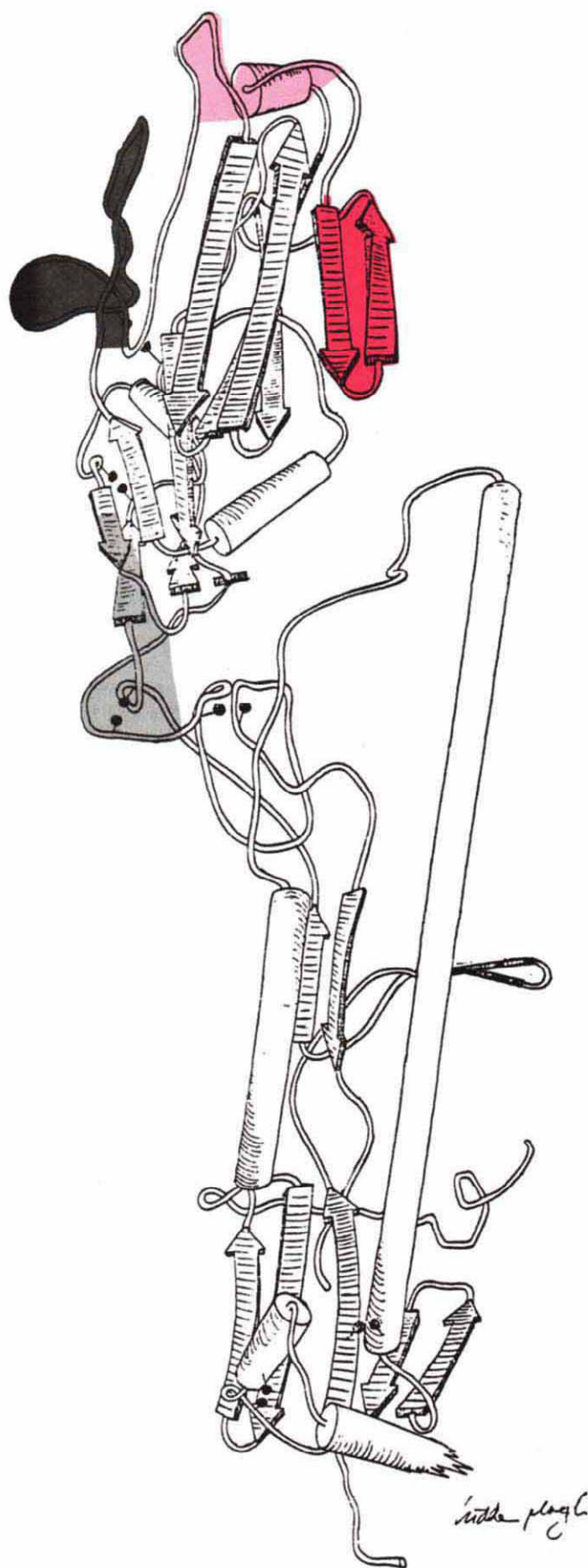


Fig. 1 Drawing (by Hidde Pleogh) of the Hong Kong (H3) HA monomer showing folding of the HA1 and HA2 polypeptides^{11,12}. The four coloured areas show where four independent antigenic areas may be located. Note that in the HA spike, which is composed of three of the monomers, the dark red site is buried and may not be involved in antibody binding.

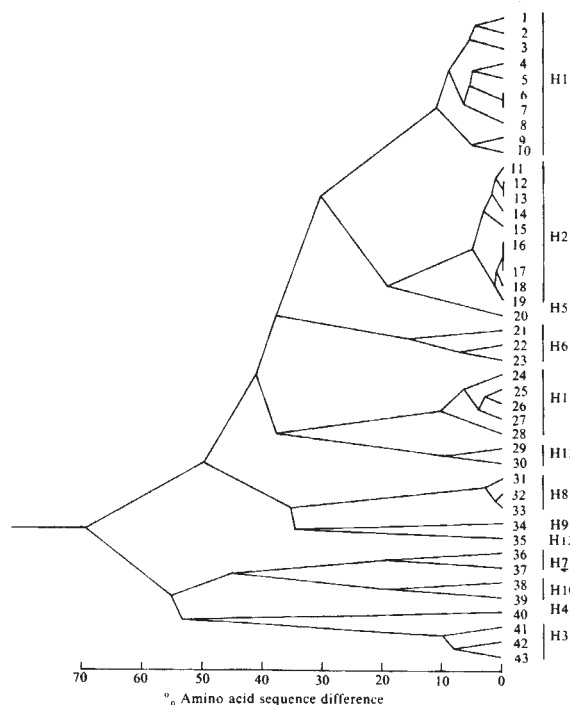


Fig. 2 Dendrogram showing the relationships between the HA1 N-terminal amino acid sequences deduced from genomic sequences of viruses representing the 13 known subtypes. The sequences were aligned using amino acids that are invariable in certain positions of all sequences (amino acid position from N terminus of most subtypes: Cys 4, Gly 6, Thr 18, Val 26, Cys 42, Cys 55, Gly 63, Pro 65, Cys 67, Glu 81). The percentage sequence differences of all pairwise comparisons of the aligned sequences starting from the N-terminal Asp or corresponding amino acid were used to calculate the dendrogram. Thus, the positions of each bifurcation in the dendrogram indicate the mean sequence difference of the sequences connected through that point⁴. The sequences used in the analysis are: 1, A/NWS/33 (H1N1); 2, A/PR/8/34 (H1N1); 3, A/BH/35 (H1N1); 4, A/Bel/42 (H1N1); 5, A/Loyang/4/57 (H1N1); 6, A/USSR/90/77 (H1N1); 7, A/Fort Warren/1/50 (H1N1); 8, A/duck/Alberta/35/76 (H1N1); 9, A/swine/Wisconsin/15/30 (H1N1); 10, A/New Jersey/11/76 (H1N1); 11, A/RI/57 (H2N2); 12, A/Tokyo/3/67 (H2N2); 13, A/Berkeley/68 (H2N2); 14, A/Netherlands/68 (H2N2); 15, A/duck/GDR/72 (H2N9); 16, A/duck/New York/6750/78 (H2N2); 17, A/duck/Ontario/77 (H2N1); 18, A/duck/Alberta/77/77 (H2N3); 19, A/pintail/Alberta/293/77 (H2N9); 20, A/shearwater/Australia/75 (H5N3); 21, A/shearwater/Australia/72 (H6N5); 22, A/duck/Alberta/1298/79 (H6N5); 23, A/duck/Alberta/1297/79 (H6N9); 24, A/tern/Australia/75 (H11N9); 25, A/duck/Ukraine/1/60 (H11N9); 26, A/duck/England/56 (H11N6); 27, A/duck/New York/12/78 (H11N6); 28, A/duck/Memphis/576/76 (H11N9); 29, A/gull/Mass/26/80 (H13N6); 30, A/gull/Md/704/77 (H13N6); 31, A/turkey/Ontario/6118/68 (H8N4); 32, A/duck/Alberta/827/78 (H8N4); 33, A/duck/Alberta/283/77 (H8N4); 34, A/turkey/Wisconsin/1/66 (H9N2); 35, A/duck/Alberta/60/76 (H12N5); 36, A/turkey/Oregon/71 (H7N3); 37, A/equine/Prague/1/56 (H7N7); 38, A/duck/Manitoba/53 (H10N7); 39, A/quail/Italy/1117/65 (H10N8); 40, A/duck/Alberta/28/76 (H4N6); 41, A/black duck/Australia/702/78 (H3N8); 42, A/Memphis/1/71 (H3N2); 43, A/duck/Ukraine/1/63 (H3N8).

stem. Each subunit has an unusual loop-like topology: it begins at the membrane, extends 135 Å distally and folds back to enter the membrane.

Four antigenic sites on the three-dimensional structure of the HA molecule of Hong Kong influenza viruses have been proposed¹²⁻¹⁴. The suggested locations (Fig. 1) were derived from amino acid sequence changes in antigenic variants of Hong Kong virus selected using monoclonal antibodies¹⁵ and in natural variants of this virus^{11,16,17}.

Variation in the HA molecule of influenza viruses occurs in two ways. It is now known that small changes in antigenicity (antigenic drift) are the result of a gradual accumulation of point mutations, while the complete change in antigenic properties (antigenic shift) involves replacement of the gene coding for one HA with that for another; this may or may not be accompanied by replacement of the NA gene.

Fig. 3 Amino acid sequences using the single letter amino acid code of the HA1 polypeptide from eight variants of Hong Kong (H3N2) influenza virus isolated between 1968 and 1979. Data are from refs 17, 21, 25, and B. A. Moss and G. Brownlee (unpublished results). Residues which have changed from NT68 are coloured or boxed. The location of most of these changes on the three-dimensional structure of the HA are shown in Fig. 1. (Boxed changes were scattered over the three-dimensional structure and may not affect the antigenic properties of the HA. They have not been shown in Fig. 1.) ● The single sequence changes found in variants selected with monoclonal antibodies¹⁵. These are (53) Asn→Lys, (133) Asn→Lys, (143) Pro→Ser, Thr, Leu, His, (144) Gly→Asp, (145) Ser→Lys, (205) Ser→Tyr. The change (226) Leu→Gln (○) appears to arise with remarkable ease. This change has been found in different clones of the same field strain and in monoclonal variants which already show another change and may be unrelated to changes in antigenicity.

	10	20	30	40	50	60
NT68	QDLPGNDNNT	ATLCLGHHAV	PNGTLVKITIT	DDQIEVTNAT	ELVQSSSTGK	ICNNPHRILD
31	QDLPGNDNST	ATLCLGHHAV	PNGTLVKITIT	DDQIEVTNAT	ELVQSSSTGK	ICNNPHRILD
NG69	QDLPGNDNST	ATLCLGHHAV	PNGTLVKITIT	NDQIEVTNAT	ELVQSSSTGK	ICNNPHRILD
U70	QDLPGNDNST	ATLCLGHHAV	PNGTLVKITIT	NDQIEVTNAT	ELVQSSSTGK	ICNNPHRILD
EM72	QDLPGNDNST	ATLCLGHHAV	PNGTLVKITIT	NDQIEVTNAT	ELVQSSSTGK	ICNNPHRILD
IC375	QDLPGNDNST	ATLCLGHHAV	PNGTLVKITIT	NDQIEVTNAT	ELVQSSSTGK	ICNNPHRILD
EX77	QNLPGNDNST	ATLCLGHHAV	PNGTLVKITIT	NDQIEVTNAT	ELVQSSSTGR	ICDSPHRILD
K79	QNLPGNDNST	ATLCLGHHAV	PNGTLVKITIT	NDQIEVTNAT	ELVQSSSTGR	ICDSPHRILD
	70	80	90	100	110	120
NT68	GIDCTLIDAL	LGDPHCDVFQ	NETWDLFVER	SKAFSNCYPY	DVPDYASLRS	LVASSGTLEF
31	GIDCTLIDAL	LGDPHCDVFQ	NETWDLFVER	SKAFSNCYPY	DVPDYASLRS	LVASSGTLEF
NG69	GIDCTLIDAL	LGDPHCDVFQ	DETWDLFVER	SKAFSNCYPY	DVPDYASLRS	LVASSGTLEF
U70	GIDCTLIDAL	LGDPHCDVFQ	NETWDLFVER	SKAFSNCYPY	DVPDYASLRS	LVASSGTLEF
EM72	GIDCTLIDAL	LGDPHCDVFQ	NETWDLFVER	SKAFSNCYPY	DVPDYASLRS	LVASSGTLEF
IC375	GIDCTLIDAL	LGDPHCDVFQ	NETWDLFVER	SKAFSNCYPY	DVPDYASLRS	LVASSGTLEF
EX77	GIDCTLIDAL	LGDPHCDVFQ	NETWDLFVER	SKAFSNCYPY	DVPDYASLRS	LVASSGTLEF
K79	GIDCTLIDAL	LGDPHCDVFQ	NETWDLFVER	SKAFSNCYPY	DVPDYASLRS	LVASSGTLEF
	130	140	150	160	170	180
NT68	ITEGFTWTGV	TQNGGSNACK	RGPSSGFFSR	LNWLTSGST	YPVLNVTMPN	NDNFDKLYIW
31	ITEGFTWTGV	TQNGGSNACK	RGPSSGFFSR	LNWLTSGST	YPVLNVTMPN	NDNFDKLYIW
NG69	ITEGFTWTGV	TQNGGSNACK	RGPSSGFFSR	LNWLTSGST	YPVLNVTMPN	NDNFDKLYIW
U70	ITEGFTWTGV	TQNGGSNACK	RGPSSGFFSR	LNWLTSGST	YPVLNVTMPN	NDNFDKLYIW
EM72	ITEGFTWTGV	TQNGGSNACK	RGPSSGFFSR	LNWLTSGST	YPVLNVTMPN	NDNFDKLYIW
IC375	ITEGFTWTGV	TQNGGSNACK	RGPSSGFFSR	LNWLTSGST	YPVLNVTMPN	NDNFDKLYIW
EX77	ITEGFTWTGV	TQNGGSNACK	RGPSSGFFSR	LNWLTSGST	YPVLNVTMPN	NDNFDKLYIW
K79	ITEGFTWTGV	TQNGGSNACK	RGPSSGFFSR	LNWLTSGST	YPVLNVTMPN	NDNFDKLYIW
	190	200	210	220	230	240
NT68	GVHHPSTNQE	QTSLYVQASG	RVTVSTRRSQ	QTIIPNIGSR	PWVRGLSSRI	SIYWTIVKPG
31	GVHHPSTNQE	QTSLYVQASG	RVTVSTRRSQ	QTIIPNIGSR	PWVRGLSSRI	SIYWTIVKPG
NG69	GVHHPSTNQE	QTSLYVQASG	RVTVSTRRSQ	QTIIPNIGSR	PWVRGLSSRI	SIYWTIVKPG
U70	GVHHPSTNQE	QTSLYVQASG	RVTVSTRRSQ	QTIIPNIGSR	PWVRGLSSRI	SIYWTIVKPG
EM72	GVHHPSTNQE	QTSLYVQASG	RVTVSTRRSQ	QTIIPNIGSR	PWVRGLSSRI	SIYWTIVKPG
IC375	GVHHPSTNQE	QTSLYVQASG	RVTVSTRRSQ	QTIIPNIGSR	PWVRGLSSRI	SIYWTIVKPG
EX77	GVHHPSTNQE	QTSLYVQASG	RVTVSTRRSQ	QTIIPNIGSR	PWVRGLSSRI	SIYWTIVKPG
K79	GVHHPSTNQE	QTSLYVQASG	RVTVSTRRSQ	QTIIPNIGSR	PWVRGLSSRI	SIYWTIVKPG
	250	260	270	280	290	300
NT68	DVLVINSNGN	LIAPRGYFKM	RTGKSSIMRS	DAPIDTCISE	CITPNGSIPN	DKPFQNVNKI
31	DVLVINSNGN	LIAPRGYFKM	RTGKSSIMRS	DAPIDTCISE	CITPNGSIPN	DKPFQNVNKI
NG69	DVLVINSNGN	LIAPRGYFKM	RTGKSSIMRS	DAPIDTCISE	CITPNGSIPN	DKPFQNVNKI
U70	DVLVINSNGN	LIAPRGYFKM	RTGKSSIMRS	DAPIDTCISE	CITPNGSIPN	DKPFQNVNKI
EM72	DVLVINSNGN	LIAPRGYFKM	RTGKSSIMRS	DAPIDTCISE	CITPNGSIPN	DKPFQNVNKI
IC375	DVLVINSNGN	LIAPRGYFKM	RTGKSSIMRS	DAPIDTCISE	CITPNGSIPN	DKPFQNVNKI
EX77	DVLVINSNGN	LIAPRGYFKM	RTGKSSIMRS	DAPIDTCISE	CITPNGSIPN	DKPFQNVNKI
K79	DVLVINSNGN	LIAPRGYFKM	RTGKSSIMRS	DAPIDTCISE	CITPNGSIPN	DKPFQNVNKI
	310	320				
NT68	TYGACPKYVK	QNTLKLATGM	RNVPEKQT			
31	TYGACPKYVK	QNTLKLATGM	RNVPEKQT			
NG69	TYGACPKYVK	QNTLKLATGM	RNVPEKQT			
U70	TYGACPKYVK	QNTLKLATGM	RNVPEKQT			
EM72	TYGACPKYVK	QNTLKLATGM	RNVPEKQT			
IC375	TYGACPKYVK	QNTLKLATGM	RNVPEKQT			
EX77	TYGACPKYVK	QNTLKLATGM	RNVPEKQT			
K79	TYGACPKYVK	QNTLKLATGM	RNVPEKQT			

Antigenic shift

Since the first human influenza virus was isolated in 1933, antigenic shifts have occurred in 1957 when the H2N2 subtype (Asian influenza) replaced the H1N1 subtype, in 1968 when the Hong Kong (H3N2) virus appeared and in 1977 when the H1N1 virus reappeared. All these major antigenic shifts in the virus occurred in China and anecdotal records suggest that previous epidemics (probably caused by the sudden emergence of 'new' viruses) also had their origin in China. Although we are still uncertain how these new subtypes appear or reappear in the human population, some clues are emerging from the plethora of recent sequence data.

The complete amino acid sequences of HA1 and HA2 have been deduced from the sequences of cloned DNA copies of the genes for strains representative of four different HA subtypes of type A influenza—fowl plague (H7N7) virus¹⁸, Asian (H2N2) influenza virus⁶, PR8 and WSN (H1N1) viruses^{19,20} and a number of strains within the human Hong Kong (H3N2) subtype^{17,21}. A complete protein sequence analysis is available only for the H3 strain^{22,23}, although sufficient protein data are available for the H2 strain to indicate where the signal peptide is removed and where the molecule is cleaved into HA1 and HA2. The data also show the location of the single disulphide bond connecting HA1 and HA2 and that most of the potential

glycosylation sites in the HA do have carbohydrate attached. The complete sequences of two strains of A/Duck/Ukraine/63 (H3N8) haemagglutinin have also been determined, one from gene sequence data²⁴, the other from direct protein sequencing²⁵.

Sequences up to 350 nucleotides from the 3' end of the HA gene and the predicted amino acid sequences from 32 type A influenza viruses, including representatives of each of the 12 known HA subtypes, have been determined²⁶. Cysteine residues and certain other amino acids are conserved in all sequences, indicating that the 12 HA subtypes evolved from a common ancestor and share a common basic structure. When the partial amino acid sequences are compared pairwise, the most distant subtypes are H1 and H3 (25% homology). Other subtypes are more similar to each other, the highest homology (80%) being between H2 and H5. Relationships between the partial sequences can be illustrated as a dendrogram (Fig. 2).

Within strains of a single subtype, the differences in deduced amino acid sequence are generally less than 10% (ref. 26), but recent extensions of the data have revealed differences in the N-terminal region of HA1 from two strains each of H7 and H10 viruses close to 20%, almost equal to the difference between the amino acid sequences of the two closest subtypes (H2 and H5).

Whereas it is possible to imagine that the H2 viruses could eventually mutate into H5, the change to H3 must involve a more drastic mechanism. How, then, could the H3 haemagglutinin have replaced the H2 strain in the human population in 1968? First, the 'new' virus may have caused an epidemic in man many years previously and have remained hidden and unchanged in some unknown place ever since. Evidence for this kind of event has been obtained. The strain of 'Russian 'flu' (H1N1) which reappeared in Anshan in northern China on 4 May 1977 and subsequently spread to the rest of the world, seems to be identical, in all genes, to the virus which caused an influenza epidemic in 1950^{27,28}. Where was this virus for 27 years? As yet we have no answer.

Second, some of the 'new' viruses may be derived from animal or avian viruses. One strain of human influenza has been shown to be such a recombinant (reassortant)²⁹⁻³¹. The Hong Kong (H3N2) virus contains the NA (and other) genes from an Asian (H2N2) strain of human influenza and a HA which is antigenically related to that of A/Duck/Ukraine/63 (H3N8) and A/equine/2/Miami/63 (H3N8) viruses^{24,25}. The amino acid sequence homology between the HAs of A/Duck/Ukraine/63 and A/Aichi/2/68 viruses (both of subtype H3) is 96%. We do not know if it was an animal or bird virus which donated the HA gene during the recombination event that led to the formation of the Hong Kong strain. The donating virus could have been one which contained Hong Kong HA that had persisted from a much earlier human influenza epidemic maintained unchanged in the same way as the Russian 'flu'. Antibodies in the sera of people who were born around 1900 suggest that a virus with a HA similar to that of the Hong Kong virus caused influenza at that time³², but the NA of this virus was N8, not N2. It is possible that viruses such as equine (H3N8) and duck/Ukraine have evolved from the 1900 epidemic strain.

The third way in which new viruses could appear in the human population is if an animal or bird virus became infectious for man. This could be brought about by mutation, but the number of mutations required and the genes in which these would need to occur are unknown.

Perhaps each of these mechanisms has operated at one time or another to produce a 'major shift' in influenza viruses infecting man. We do not know how the different subtypes of type A influenza arose, but it is reasonable to suppose that 'shuffling' of genes between influenza A strains is occurring all the time, most frequently in birds³³, but also in mammals, including man^{34,35}.

Antigenic drift

Antigenic drift in the HA occurs by mutations in the gene, leading to an accumulation of amino acid sequence changes that alter the antigenic sites in such a way that they are no longer

recognized by the host's immune system. The evidence for this mechanism comes from sequence analysis of naturally occurring 'drifted' viruses (mainly those of the Hong Kong (H3N2) virus) and of variants selected in the laboratory in the presence of antibodies. Most of the sequence changes occurring in Hong Kong HA between 1968 and 1979 were in the HA1 portion of the HA molecule, only three changes being found in HA2¹⁷. The sequence changes in HA1 during this period are shown in chronological order of isolation of the strains in Fig. 3. Although these changes are distributed over HA1, on the three-dimensional structure of the HA many are clustered into four regions (Fig. 1) which are assumed to be the antigenic sites on the HA. Sequential changes at a particular position in HA1 have not occurred. A few apparent reversions were seen, but as there is no reason to believe that the viruses examined represent a direct genealogical lineage, it is likely that in those cases in Fig. 3 where there are apparent reversions of an amino acid, the real explanation is that the residue never changed in the first place. Similarly, in the one case where there is an apparent sequential change in an amino acid (the asparagine residue at position 137 changing to serine in Vic/75 and then to tyrosine in Bangkok/1/79), the nucleotide changes involved indicate that the Vic/75 strain was not the progenitor of the Bangkok/1/79 strain.

Some regions of HA1 have been totally conserved during the period 1968-79. The largest of these conserved regions occurs between residues 84-121 and 279-328. The latter block is also conserved in the two avian H3 strains so far sequenced^{24,25}, except for a single change of valine to isoleucine at position 309. Residues 279-328 form part of the 'stalk' of the HA where structural constraints may necessitate that they are highly conserved.

Variants selected with monoclonal antibodies

Antigenic variants of the HA have been selected by growing virus in the presence of monoclonal antibody to the HA. The variants occurred at a frequency of $\sim 10^{-5}$ in virus stock^{36,37} and did not bind at all to the antibody used for their selection. This dramatic change in antigenicity was found to be associated with single changes in the amino acid sequence of HA1^{15,37} (see Fig. 3). Most antigenic variants of Mem/1/71 (H3N2) selected with monoclonal antibodies could not be distinguished from parental virus with antisera from ferrets infected with the parent strain, but two could be distinguished¹³. Some antigenic variants of B/Hong Kong/1/73 could also be discriminated³⁸. Presumably, their amino acid substitutions (Fig. 3) cause more profound changes in the antigenicity of the molecule than others. It seems likely that only variants which have such profound changes have sufficient selective survival advantages in an immune population to produce an epidemiological impact.

Variation in neuraminidase

Neuraminidase is an integral membrane protein in influenza virus and can be solubilized with detergents or pronase³⁹. The detergent-soluble NA has a molecular weight^{40,41} of $\sim 240,000$ and consists of four identical polypeptide chains arranged in such a way that the molecule possesses a hydrophilic 'head' attached to the hydrophobic membrane-insertion sequence by a thin stalk. The active centre of the enzyme and the antigenic determinants are distinguishable from each other⁴². Pronase digestion of the virus releases the soluble NA head of $M_r \sim 200,000$ which retains the antigenic and enzymatic activity of the intact molecule. The NA heads of some strains of influenza virus can be crystallized⁴³ and determination of the three-dimensional structure of A/Tok/3/67 (H2N2) NA is in progress⁴⁴. The Fab fragment from a monoclonal antibody to A/Tok/3/67 (H2N2) NA has also recently been crystallized and the complementary surfaces of this antigen-antibody complex can now be mapped⁴⁵.

The complete sequence of the NA gene of A/PR/8/34 has been determined⁴¹ and from the predicted amino acid sequence and analysis of intact NA and NA heads of N2, N5 and N8

Table 1 Influenza virus-coded proteins

Designation	Approximate no. of molecules per virus particle	Molecular weight estimated by:			Ref.	Remarks
		Gel electrophoresis	Gene sequence*	Gene + protein sequence		
P1 polymerase 1	30-60	96,000			91	Internal proteins associated with RNA transcriptase activity.
P2 polymerase 2		87,000			91	
P3 polymerase 3		85,000			91	
HA haemagglutinin	500					
HA1				36,074 + 11,500 (Mem/102/72)†	22,84	Surface glycoprotein responsible for attachment of virus to cells.
HA2				27,368 + 1,400 (Mem/102/72)†	23,84	Trimer composed of two polypeptides HA1 and HA2 formed by post-translational cleavage of the primary translation product. Three-dimensional structure known.
NP nucleoprotein	1,000	50-60,000	56,106		55	Internal protein associated with RNA and polymerase proteins, helical arrangement.
NA neuraminidase	100	48-63,000	50,087		41	Surface glycoprotein, with enzyme activity. Tetragonal molecule with a 200,000- M_r head—crystallized and structure in progress.
M1 matrix	3,000			27,861	58, 59, 61	Major virion component surrounding the core, involved in assembly and budding.
M2 matrix		15,000	11,000		10	Coded from the same gene segment as M1 in a second reading frame, a non-structural protein, function unknown.
NS1 non-structural protein		25,000	26,815		64	Non-structural protein, function unknown.
NS2 non-structural protein		12,000	14,216		64	Non-structural protein coded for in a second reading frame from same gene segment as NS1, function unknown, synthesized late in infection.

* These values are from the predicted amino acid sequence and do not take into account any processing the final product may have undergone.

† The second value is the carbohydrate contribution.

subtypes^{7,46} it has been found that, unlike the HA, the NA is oriented with its N terminus buried in the viral membrane and does not have a signal peptide which is cleaved.

Antigenic shift and antigenic drift occur in the NA; among the human influenza viruses antigenic shift occurred in 1957, with the emergence of the H2N2 subtype and the replacement of the N1 by N2. Antigenic drift in the NA of influenza viruses⁴⁷⁻⁴⁹ has been correlated with differences in amino acid sequences^{43,50}.

Partial sequences up to 340 nucleotides from the 3' end of the NA gene have been obtained for representative viruses from eight of the nine NA subtypes⁵¹. These sequences will correspond to the region encoding the N-terminal part of the NA since analysis of tryptic peptides shows that translation begins from the first AUG codon and that no signal peptide is cleaved from the N terminus. The first six amino acids are conserved in all viruses of all subtypes examined and the next six are conserved in most subtypes. Following these, the nucleotide and predicted amino acid sequences of the eight subtypes of NA differ dramatically in the hydrophobic membrane insertion sequence and the extended structure of the stalk. Deletions and/or insertions of short blocks of nucleotides in the NA gene are seen within subtypes⁵².

Antigenic variants of the NA have been selected using monoclonal antibodies at a frequency of ~ 1 in 10^5 (ref. 53), similar to the frequency of variants in the HA molecule³⁶. Analysis of antigenic variants with a panel of N2 monoclonal antibodies provided evidence for at least three and possibly four non-overlapping antigenic areas on the NA molecule of A/Tokyo/3/67⁵³. Studies on natural antigenic drift indicate that it occurs by the same mechanisms as in the HA, with a limited number of base changes in the NA gene within subtypes⁵⁴.

Variation in the nucleoprotein

The nucleoprotein (NP) is one of the group-specific antigens of influenza viruses that distinguishes between the influenza A, B and C viruses. It probably constitutes the backbone of the helical internal complex that is associated with the RNA segments and the three different polymerases.

The NP gene of A/PR/8/34 virus is 1,565 nucleotides long and is capable of encoding a protein of 498 amino acid residues

(M_r 56,106) rich in arginine⁵⁵. Double immunodiffusion tests showed antigenic differences between the NPs of H1N1 and the H3N2 strains⁵⁶ and recent studies with monoclonal antibodies to the NP of A/WSN/33 (H1N1) viruses have shown that antigenic variation occurs in this molecule⁵⁷. The NP molecule possesses at least three non-overlapping antigenic areas, one area being the same on all strains tested. Monoclonal antibodies to this conserved domain inhibited *in vitro* transcription of viral RNA, suggesting that this region of the NP is involved in RNA transcription⁵⁷.

The matrix protein

The complete sequence of RNA segment 7 (M) of two strains of A/PR8/34 (H1N1)^{58,59} and of A/Udorn/72 (H3N2)⁶⁰ has been reported as well as partial sequences of a number of strains⁶¹. Following the first AUG codon in the positive strand, a 252-residue protein, rather hydrophobic and rich in arginine, is encoded. Compositions of tryptic peptides from purified matrix protein of PR8⁶¹ correspond well with those expected from the amino acid sequence predicted from the gene sequence.

Comparison of the sequences of RNA segment 7 of the H3N2 (Udorn) and H1N1 (PR8) strains shows that the sequences coding for the matrix (M) of these viruses isolated 38 years apart are highly conserved⁶⁰, in keeping with antigenic studies⁶². Comparison of 230 nucleotides of RNA segment 7 from five human H1N1, H2N2 and H3N2 strains isolated over a 43-yr period suggests that the same segment 7 was retained throughout the antigenic shifts of HA and NA type (H1N1 to H2N2 to N3N2)⁶¹. In addition, the complete sequences contain a second open reading frame which overlaps the M protein sequence by 68 nucleotides.

Three mRNAs transcribed from RNA segment 7 have been isolated. One (M1 mRNA) consists of an uninterrupted, nearly full-length, copy of RNA segment 7 and is responsible for the production of the M protein. An M2 protein is generated from a spliced product of the M1 mRNA, such that after the nucleotides encoding the N-terminal nine amino acids, nearly 600 nucleotides are spliced out and the reading frame is changed; a protein product corresponding to this has been identified in infected cells¹⁰. In addition, a third mRNA has been found, which would code only for a 8-residue peptide identical to the C terminus of M1⁶³. Such a product has not yet been isolated.

The non-structural proteins

Recent studies have shown that RNA segment 8 codes for at least two non-structural polypeptides, NS1 and NS2, which are translated from separate mRNAs⁶⁴⁻⁶⁶. Mapping and sequence studies have shown that NS1 and NS2 overlap by 70 amino acids that are translated from different reading frames. Polypeptides NS1 and NS2 share 9 amino acids at their N termini, but after this sequence the mRNA for NS2 has a deletion of 423 nucleotides, then rejoins the rest of the mRNA in the +1 reading frame. The function of NS1 or NS2 has not been established. NS1 is made in large amounts and accumulates in the nucleus⁶⁷; NS2 is made late in infection and is found predominantly in the cytoplasm^{64,68}.

Because NS1 and NS2 are internal proteins of infected cells and hence less available to antibodies, they would be expected to show less sequence variation than the surface glycoproteins (HA and NA). Accordingly, comparison of the sequences of the NS genes of fowl plague⁶⁵ and the two human influenza strains A/Udorn/72 (H3N2) and A/PR/8/34 shows only 8-11% differences^{64,69,70}. An open reading frame potentially coding for a polypeptide has been noted in the noncoding, virion RNA of the NS genes of A/PR/8/34, Udorn/72 and FPV, but is not present in the NS gene of duck/Alberta/60/76⁷¹. No protein corresponding to this extra 'gene' has yet been identified.

Polymerase proteins

The three largest proteins of the virion (P1, P2, P3) with M_s of 96,000, 87,000 and 85,000, respectively, are found in association with the nucleoprotein and virion RNA and carry the polymerase activity⁷² which transcribes the invading viral RNA⁷³. Proteins P1 and P3 are probably required for complementary RNA synthesis and P2 and NP for virion RNA synthesis⁷⁴.

Transcription of the viral RNA segments soon after infection terminates ~40 nucleotides before the 5' end to give messenger RNAs⁷⁵; these mRNAs have a cap structure and a heterogeneous sequence of 10-13 nucleotides derived from cellular mRNAs at their 5' end⁷⁶. Later transcripts are full length and act as templates for synthesis of the genes of the progeny virus. The complete nucleotide sequences of two of the three polymerase genes of the A/PR/8/34 strain have been determined⁷⁷, but the extent of variation in the polymerase genes is unknown, although these may play an important part in host range and virulence.

Defective interfering (DI) influenza virus particles are generated by high multiplicity passage in permissive cells. These particles are of interest because they facilitate the establishment of persistent infection in cell cultures and could therefore be concerned in latency. They contain new small RNA molecules which are absent from infectious virus and which are generated predominantly by massive internal deletion from the three P genes^{78,79}, although the sequences of one small RNA is a mosaic of several segments from at least two of the polymerase genes^{77,80}. This is the first time intragenic recombination has been shown in influenza virus.

Variation in pathogenicity

Highly pathogenic avian influenza viruses that cause generalized infections and death in avian species exist among viruses of the H7 (formerly H_{av}1) and H5 (H_{av}5) HA subtypes, while other viruses of these subtypes are relatively non-pathogenic although fully infectious. Efficient cleavage of the precursor HA polypeptide into HA1 and HA2 in a wide range of cells from different tissues appears to be essential for the pathogenicity of these viruses⁸¹⁻⁸³. The amino acid sequence of the connecting peptide between HA1 and HA2 may determine the range of cells in which cleavage can occur. Fowl plague virus, which produces almost invariably lethal infections in chickens, has the sequence Lys-Lys-Arg-Gln-Lys-Arg between the C terminus of HA1 and the N terminus of HA2¹⁸, while the human viruses, which are avirulent in chickens, and non-pathogenic avian viruses of the same subtype as fowl plague (H7), contain only a

single connecting arginine residue^{6,17,22,23,84}. Investigation of groups of pathogenic and non-pathogenic avian influenza viruses of the H7 subtype suggested that the HA molecules from the non-pathogenic viruses had fewer basic residues in their connecting HA peptides than the pathogenic strains⁸⁵.

Examination of the amino acids of the amino-terminal sequence of HA2 of several influenza A viruses of human and avian origin has revealed a highly conserved region of 10 residues^{86,87}. This sequence is homologous⁸⁸ to the amino-terminus region of the fusion glycoprotein F, of Sendai virus (a paramyxovirus), which mediates fusion of the Sendai virus envelope with the plasma membrane of host cells, a function which is considered essential for virus infectivity. It has been postulated, therefore, that the infectivity of influenza virus is dependent on a fusion function of HA analogous to that of the Sendai fusion protein. Although the three-dimensional structure of the influenza HA shows that this peptide is buried in the molecule, it has recently been shown that after incubation in conditions required for membrane fusion, that is, at pH 5.0, a conformational change occurs in the influenza HA glycoprotein and the molecule acquires the ability to bind to lipid vesicles⁸⁹.

However, the structure of the HA is almost certainly not the only factor determining pathogenicity. While reassortment studies between highly virulent and avirulent viruses pointed to the prime importance of the HA for pathogenicity, they also showed that the polymerase (P3) and nucleoprotein genes could be involved⁹⁰. Other experiments have shown that a specific gene constellation, rather than any single gene, is responsible for the pathogenic properties⁹¹.

What of the future?

The recent explosion of structural data on the genes and gene products of influenza viruses has provided some important information about the behaviour of this virus. For example, it is clear that antigenic shift does not occur by direct mutation of one subtype into another, while antigenic drift occurs by point mutation in the genes. However, the mutation frequency of the surface proteins is not significantly greater than that of the internal proteins. HA possesses at least four distinct antigenic sites which can vary independently, with pathogenicity depending in part on the sequence in the HA where cleavage occurs. Some RNA segments of the virus possess overlapping genes, using two reading frames. The sequence data support the antigenic classification of the HA and NA subtypes.

However, many mysteries remain. Shifts can occur by the emergence of viruses which previously caused epidemics and have remained unchanged, as if frozen, for many years. Where these viruses hide, and what causes them to re-emerge, is a complete mystery. The shifts may also occur by genetic reassortment of human and animal influenza viruses or by mutation of an animal or avian virus so that it becomes able to infect people. Most type A influenza virus subtypes are found in birds, but the role of these in the emergence (or re-emergence) of new subtypes in man is not clear; nor is it clear why the shifts always occur in China or why the old subtypes usually disappear from man as the pandemic of new virus starts. On the other hand, the Hong Kong subtype did not disappear when Russian 'flu started and we do not know why this was.

What is the origin of the various subtypes? How many subtypes are there in birds and animals as well as in man? Can all subtypes of HA and NA be incorporated into viruses infecting man? Why are the antigenic shifts restricted to type A influenza? What triggers the appearance of a new subtype?

It is not known precisely how the amino acid sequence changes which occur in the HA and NA during drift alter the antigenic properties. Does drift follow a set path and can drift be predicted? Why do few, if any, sequential changes at particular positions in the HA molecule occur?

Are there particular areas on the HA and NA which induce neutralizing antibody while others do not? Does drift involve changes in the immunogenicity of different antigenic determinants? The location on the HA molecule of the sequence

changes which occur during antigenic variation in the H3 subtype is known, but whether the amino acids in these areas make contact with the combining site on antibody molecules is not.

Does drift continue indefinitely? Can it go backwards? It is not known if the same antigenic change occurs simultaneously in different areas of the world or whether a single mutant arises somewhere and all spread derives from a single focus. Why does antigenic drift occur in influenza, but not, for example, in paramyxoviruses such as measles virus, where antigenic variants can be selected *in vitro* with monoclonal antibodies at about the same frequency as influenza^{92,93}?

Why does one sometimes observe 'original antigenic sin', in which a virus may preferentially elicit antibodies against a previous strain which has since disappeared? In addition to antigenic variation, variation in host range, virulence and transmissibility greatly influence the severity of the disease and the spread of epidemics. Is it possible that a virus could appear which would cause the same degree of mortality as the 1918 'swine' flu epidemic? What influences the spread of influenza virus, and why do epidemics normally appear in the winter, thus oscillating between the Northern and Southern Hemispheres? It is not known why epidemics are self-limiting, even when susceptible hosts remain in abundance.

There are now, however, better prospects for the control of influenza. Effective and safe live vaccines may be made in the future using modern genetic engineering techniques. Synthetic vaccines may be possible where the antigens are pure enough for the vaccines to be given in large enough doses to be effective and engineered such that they do not induce original antigenic sin. A universal vaccine, which protects against all members of a

subtype or even across subtypes might be feasible. It might also be more rewarding to look for ways of increasing the cellular immune response as well as the humoral response.

With the three-dimensional structure of at least one HA molecule known and candidates for the antigen and receptor binding sites tentatively identified, the way is open for the synthesis of vaccines containing key segments of HA molecules. HA antigenic determinants can be detected in *Escherichia coli* transformed with a plasmid carrying a cloned HA gene⁹⁴⁻⁹⁶, and copious amounts of HA appear on the surfaces of eukaryotic cells infected with a simian virus 40 (SV40) vector containing the HA gene^{97,98}. These experiments open the way to the use of site-directed mutagenesis to tailor the HA into a successful synthetic and possibly universal vaccine.

What of the immediate future? We do not know if another antigenic shift is imminent. If a new virus does suddenly appear, will the structural data tell us where it came from? The new virus might not have the capacity to kill people in the same way as the recent seal virus killed seals, but if it did, it is doubtful that we would be able to stop it causing world-wide devastation. If no shift occurs, what will happen to the currently circulating H3N2 and H1N1 strains? Will they continue to drift, or have they reached their limit? There is little doubt that a greater understanding of the molecular biology and genetics of the virus will make an essential contribution in piecing together the answers to the puzzle of influenza.

The authors were supported in part by grants AI08831, AI02649 and AI15343 from the National Institute of Allergy and Infectious Diseases. The writing of this joint review was greatly assisted by international telephone facilities donated by the Australian Overseas Telecommunications Commission.

- Geraci, J. R. *et al.* *Science* (in the press).
- Webster, R. G. *et al.* *Virology* **113**, 712-724 (1981).
- World Health Organization Memorandum *Bull. Wld Hlth Org.* **58**, 585-591 (1980).
- Hinshaw, V. S. *et al.* *J. Virol.* (in the press).
- Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 93-97 (1975).
- Gething, M. J., Bye, J., Skehel, J. J. & Waterfield, M. *Nature* **287**, 301-306 (1980).
- Blok, J. *et al.* *Virology* (in the press).
- Lazarowitz, S. G., Compans, R. W. & Choppin, P. W. *Virology* **46**, 830-843 (1971).
- Skehel, J. J. *Virology* **49**, 23-36 (1972).
- Lamb, R. & Choppin, P. W. *Virology* **112**, 729-737 (1981).
- Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366-373 (1981).
- Wiley, D. C., Wilson, I. A. & Skehel, J. J. *Nature* **289**, 373-378 (1981).
- Webster, R. G. & Laver, W. G. *Virology* **104**, 139-148 (1980).
- Gerhard, W., Yewdell, J. & Frankel, M. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 273-280 (Elsevier, New York, 1980).
- Laver, W. G., Air, G. M. & Webster, R. G. *J. molec. Biol.* **145**, 339-361 (1981).
- Laver, W. G., Air, G. M., Dopheide, T. A. & Ward, C. W. *Nature* **283**, 454-457 (1980).
- Both, G. W. & Sleight, M. J. *J. Virol.* **39**, 663-672 (1981).
- Porter, A. G. *et al.* *Nature* **282**, 471-477 (1979).
- Winter, G., Fields, S. & Brownlee, G. G. *Nature* **292**, 72-75 (1981).
- Hitti, A. L., Davis, A. R. & Nayak, D. P. *Virology* **111**, 113-124 (1981).
- Verhoeyen, M. *et al.* *Nature* **286**, 771-776 (1980).
- Ward, C. W. & Dopheide, T. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 27-38 (Elsevier, New York, 1980).
- Dopheide, T. A. & Ward, C. W. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 21-26 (Elsevier, Amsterdam, 1980).
- Fang, R., Min Jou, W., Huylebroeck, D., Devos, R. & Fiers, W. *Cell* **25**, 315-323 (1981).
- Ward, C. W. & Dopheide, T. *ICN-UCLA Symp. molec. cell. Biol.* **22**, 331-340 (1982).
- Air, G. M. *Proc. natn. Acad. Sci. U.S.A.* **58**, 7639-7643 (1981).
- Nakajima, K., Desselberger, U. & Palese, P. *Nature* **274**, 334-339 (1978).
- Scholtissek, C., van Hovningen, V. & Rott, R. *Virology* **89**, 613-617 (1978).
- Laver, W. G. & Webster, R. G. *Br. med. Bull.* **35**, 29-33 (1979).
- Laver, W. G. & Webster, R. G. *Virology* **51**, 383-391 (1973).
- Scholtissek, C., Rohde, W., Harms, E. & Rott, R. in *Negative Strand Viruses and the Host Cell* (eds Barry, R. D. & Mahy, B. W. J.) (Academic, London, 1978).
- Masurel, N. *Bull. Wld Hlth Org.* **41**, 461-468 (1969).
- Hinshaw, V. S., Webster, R. G., Bean, W. J. & Sriam, G. *Comp. Immun. Microbiol. infect. Dis.* **3**, 155-164 (1980).
- Young, J. F. & Palese, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6547-6551 (1979).
- Bean, W. J., Cox, N. J. & Kendal, A. P. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 105-114 (Elsevier, New York, 1980).
- Yewdell, J. W., Webster, R. G. & Gerhard, W. *Nature* **279**, 246-248 (1979).
- Laver, W. G. *et al.* *Virology* **98**, 226-237 (1979).
- Webster, R. G. & Berton, M. T. *J. gen. Virol.* **54**, 243-251 (1981).
- Drzeniek, R. *Curr. Topics Microbiol. Immun.* **59**, 35-74 (1972).
- Wrigley, N. G. *Br. med. Bull.* **35**, 35-38 (1979).
- Fields, S., Winter, G. & Brownlee, G. G. *Nature* **290**, 213-217 (1981).
- Fazekas de St. Groth, S. *Ann. N.Y. Acad. Sci.* **103**, 674-685 (1963).
- Laver, W. G. *Virology* **86**, 78-87 (1978).
- Colman, P. M. & Laver, W. G. in *Proc. 7th Aharon-Katzir-Katchalsky Conf. Structural Aspects of Recognition and Assembly in Biological Macromolecules* (eds Balaban, M., Sussman, J. L., Traub, W. & Yonath, A.) 869-872 (Balaban ISS, Rehovot, 1981).
- Colman, P. M. *et al.* *J. molec. Biol.* **152**, 609-614 (1981).
- Blok, J. & Air, G. M. *Biochemistry* (submitted).
- Paniker, K. J. *J. gen. Virol.* **2**, 385-394 (1968).
- Schild, G. C. & Newman, R. W. *Bull. Wld Hlth Org.* **41**, 437-445 (1969).
- Curry, R. L., Brown, J. D., Baker, F. A. & Hobson, D. J. *J. Hyg.* **72**, 197-204 (1974).
- Kendal, A. P. & Kiley, M. P. *J. Virol.* **12**, 1482-1490 (1973).
- Blok, J. *ICN-UCLA Symp. molec. cell. Biol.* **22**, 45-54 (1982).
- Blok, J. & Air, G. M. *Virology* (in the press).
- Webster, R. G., Hinshaw, V. S. & Laver, W. G. *Virology* (in the press).
- Blok, J. & Air, G. M. *Virology* **107**, 50-60 (1980).
- Winter, G. & Fields, S. *Virology* **114**, 423-428 (1981).
- Schild, G. C., Oxford, J. S. & Newman, R. W. *Virology* **93**, 569-573 (1979).
- Van Wyke, K. L., Hinshaw, V. S., Bean, W. J. & Webster, R. G. *J. Virol.* **35**, 24-30 (1980).
- Winter, G. & Fields, S. *Nucleic Acids Res.* **8**, 1965-1974 (1980).
- Allen, H., McCauley, J., Waterfield, M. J. & Gething, M.-J. *Virology* **107**, 548-551 (1980).
- Lamb, R. A. & Lai, C. J. *Virology* **112**, 746-751 (1981).
- Hall, R. M. & Air, G. M. *J. Virol.* **38**, 1-7 (1981).
- Schild, G. C. *J. gen. Virol.* **15**, 99-103 (1972).
- Lamb, R. A., Lai, C. J. & Choppin, P. W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4170-4174 (1981).
- Lamb, R. & Lai, C. J. *Cell* **21**, 475-485 (1980).
- Porter, A. G., Smith, J. C. & Emtage, J. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5074-5078 (1980).
- Inglis, S. C., Barrett, T., Brown, C. M. & Almond, J. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3790-3794 (1979).
- Lazarowitz, S. G., Compans, R. W. & Choppin, P. W. *Virology* **46**, 830-843 (1971).
- Mahy, B. W. J., Barrett, T., Briedis, D. J., Brownson, J. M. & Wolstenholme, A. J. *Phil. Trans. R. Soc. B288*, 349-357 (1980).
- Baez, M., Taussig, R., Zazra, J. J., Young, J. F. & Palese, P. *Nucleic Acids Res.* **8**, 5845-5858 (1980).
- Winter, G., Fields, S., Gait, M. & Brownlee, G. *Nucleic Acids Res.* **9**, 237-245 (1981).
- Baez, M., Zazra, J. J., Elliott, R. M., Young, J. F. & Palese, P. *Virology* **113**, 397-402 (1981).
- Bishop, D. H., Roy, P., Bean, W. J. & Simpson, R. W. *J. Virol.* **10**, 689-697 (1972).
- Bean, W. J. & Simpson, R. W. *Virology* **56**, 646-651 (1973).
- Palese, P. & Ritchey, M. B. *Virology* **78**, 183-191 (1977).
- Hay, A. J. & Skehel, J. J. *Br. med. Bull.* **35**, 47-50 (1979).
- Plotch, S. J., Bouloy, M. & Krug, R. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1618-1622 (1979).
- Fields, S. & Winter, G. *Cell* (in the press).
- Nayak, D. P. A. *Rev. Microbiol.* **34**, 619-644 (1980).
- Nakajima, K., Ueda, M. & Sugiura, A. *J. Virol.* **29**, 1142-1148 (1979).
- Moss, B. A. & Brownlee, G. *Nucleic Acids Res.* **9**, 1941-1947 (1981).
- Klenk, H.-D., Rott, R., Orlich, M. & Blödmann, J. *Virology* **68**, 426-439 (1975).
- Lazarowitz, S. G. & Choppin, P. W. *Virology* **68**, 440-454 (1975).
- Rott, R., Reinacher, M., Orlich, M. & Klenk, H.-D. *Arch. Virol.* **65**, 123-133 (1980).
- Both, G. W. & Sleight, M. J. *Nucleic Acids Res.* **8**, 2561-2575 (1980).
- Bosch, F. X., Garten, W., Klenk, H.-D. & Rott, R. *Virology* **113**, 725-735 (1981).
- Waterfield, M. D., Espelie, K., Elder, K. & Skehel, J. J. *Br. med. Bull.* **35**, 57-63 (1979).
- Scheid, A., Graves, M. C., Silver, S. M. & Choppin, P. W. in *Negative Strand Viruses and the Host Cell* (eds Mahy, B. W. J. & Barry, R. D.) 181-193 (Academic, London, 1978).
- Gething, M.-J., White, J. M. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2737-2740 (1978).
- Skehel, J. J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Bean, W. & Webster, R. G. *Negative Strand Viruses and the Host Cell* (eds Mahy, B. W. J. & Barry, R. D.) 685-692 (Academic, London, 1978).
- Scholtissek, C. *Curr. Topics Microbiol. Immun.* **80**, 139-169 (1978).
- Portner, A., Webster, R. G. & Bean, W. J. *Virology* **104**, 235-238 (1980).
- Birrer, M. J., Udern, S., Nathenson, S. & Bloom, B. R. *Nature* **293**, 67-69 (1981).
- Emtage, J. S. *et al.* *Nature* **283**, 171-174 (1980).
- Heiland, I. & Gething, M.-J. *Nature* **292**, 851-852 (1981).
- Davis, A. R. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 5376-5380 (1981).
- Gething, M.-J. & Sambrook, J. *Nature* **293**, 620-625 (1981).
- Sveda, M. M. & Lai, C.-J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5488-5492 (1981).

ARTICLES

Revisions of K/Ar ages for the Hadar hominid site, Ethiopia

Robert C. Walter* & James L. Aronson

Department of Geological Sciences, Case Western Reserve University, Cleveland, Ohio 44106, USA

Conventional K/Ar age measurements are reported for two volcanic units from the hominid-bearing Hadar Formation. These ages represent revision of previously published ages for the BKT-2 tephra and the Kadada Moumou basalt. BKT-2 is revised from 2.65 to 2.9 Myr BP and the Kadada Moumou basalt is revised from 3.0 to 3.6 Myr BP.

PREVIOUS isotopic age determinations for the Hadar Formation relied on two volcanic horizons, the BKT-2 tephra in the upper Kada Hadar (KH) Member and the Kadada Moumou basalt (KMB) in the upper Sidi Hakoma (SH) Member (Fig. 1). New work has focused on additional dating of these two units as well as on many of the other tuffs in the formation that are less adequate for dating¹. This article presents revised geochronology for the BKT-2 tuff and of the Kadada Moumou basalt. The revisions are based on three components: (1) new abundance and decay constants for ⁴⁰K that are now widely accepted²; (2) new values for the ³⁸Ar internal standard used in the study; and (3) running of additional samples that have shed new light on the age interpretation. These considerations cause us to propose an increased age for these units.

New decay and abundance constants for ⁴⁰K

In August 1977, the International Union of Geological Sciences (IUGS) Subcommittee on Geochronology proposed a revision of the decay constants and the atomic abundance of ⁴⁰K (ref. 2). The new constants are $\lambda_e = 0.581 \times 10^{-10} \text{ yr}^{-1}$, $\lambda_\beta = 4.962 \times 10^{-10} \text{ yr}^{-1}$, and $^{40}\text{K}/\text{K} = 1.167 \times 10^{-4}$. Plio-Pleistocene K/Ar ages calculated with the old constants are increased by ~2.7% when recalculated with the new constants³. Therefore, previously published K/Ar ages for the Pliocene Hadar Formation⁴⁻⁶ must be categorically increased by 2.7%.

Revision of CWRU ³⁸Ar pipette constants

The K/Ar laboratory at Case Western Reserve University (CWRU) utilizes a bulb ³⁸Ar tracer system with an all-metal, two-valve gas pipette machined in Heidelberg after the type used at the Max Planck Institute. The bulb was filled with the ³⁸Ar residue remaining after a commercial preparation of ³⁸Ar batch tracers. A discrepancy arose between the pipette volume calculated from the amount of ³⁸Ar stipulated for the initial fill by the commercial glassblower and the pipette volume measured by the machinist who made the pipette. In our previous reports on the Hadar Formation⁴⁻⁶ the 'calculated' rather than the 'measured' volume was used and we reported an average of $18.4 \times 10^{-10} \text{ mol g}^{-1}$ of ⁴⁰Ar* (radiogenic argon) for the LP6 interlaboratory biotite standard. This value is significantly lower than the accepted LP6 ⁴⁰Ar* yield of $19.3 \times 10^{-10} \text{ mol g}^{-1}$ (J. C. Engles, personal communication).

We have now pipetted over 1,000 aliquots from the bulb system and monitored its depletion against 35 analyses of the LP6 standard (Table 1). The depletion behaviour (Fig. 2) is only explained by the machinist's measurement of the pipette volume (method B) and not by the glassblower's calculation of the initial ³⁸Ar content (method A). Thus we have recalibrated our ³⁸Ar internal standard to the measured pipette volume and recalculated the initial ³⁸Ar fill to correspond to the accepted value of $19.3 \times 10^{-10} \text{ mol g}^{-1}$ of ⁴⁰Ar* for the LP6 biotite standard.

Table 1 shows our results for several inter-laboratory standards including LP6 biotite, MMhb hornblende (University of Minnesota), SB biotite (US Geological Survey), Bern 4M muscovite and Bern 4B biotite and Berkeley African Feldspar: (BAFIS) A, B and C (ref. 7 and personal communications from J. C. Engles, M. Lanphere, and E. Jager).

In Tables 2 and 5 all previously published K/Ar ages for the Hadar Formation have been corrected for the recalibrated ³⁸Ar pipette constant as well as the new radiometric and abundance constants.

Additional samples and new measurements

BKT-2, believed to be a primary ash-fall tuff, overlies nearly all of the fossil vertebrate fauna of the Hadar Formation⁵ (Fig. 1). The tuff commonly occurs as a doublet horizon, with the upper ash layer (BKT-2u) separated from the lower ash layer (BKT-2L) by 0.5–1.0 m of fluvial overbank silty clay⁸.

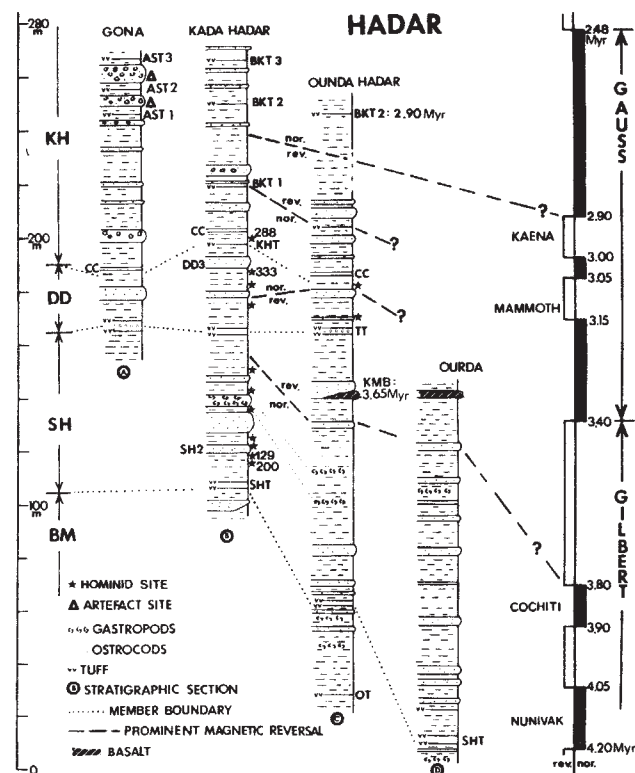


Fig. 1 Stratigraphical sections proceeding from west (left) to east (right) across the Hadar Site (taken from ref. 8). Location of sections shown in Fig. 3. Prominent magnetic polarity reversals and their possible correlation to the worldwide reversal time scale (ref. 3) are indicated.

* Present address: Department of Physics/Geophysics, University of Toronto, Toronto, Ontario, Canada M5S 1A7.

Table 1 Interlaboratory K/Ar standards

Sample	% K ₂ O		⁴⁰ Ar* (10 ⁻¹⁰ mol g ⁻¹)		No. of measurements
	Measured	Accepted	Measured	Accepted	
LP6 biotite†			19.25±0.56	(19.30)	35*
MMhb hornblende ⁷	1.89	1.874	16.31±0.10	(16.24±0.09)	4
SB biotite‡	9.188	9.188	22.78±0.11	(22.45±0.29)	3
Bern 4M muscovite§	10.58	10.48	2.716±0.015	(2.82)	4
Bern 4B biotite§	9.57	9.52	2.446±0.014	(2.38)	2
Berkeley African feldspar A	4.725	(¶)	0.1233 (1.81 Myr BP)	(¶)	1
Berkeley African feldspar B					
run 1	4.49	(¶)	0.1054 (1.63 Myr BP)	()	1
run 2			0.0941 (1.46 Myr BP)	()	1
Berkeley African feldspar C	5.74	(¶)	0.1531 (1.85 Myr BP)	()	1

Personal communications from † J. C. Engles, ‡ M. Lanphere, § E. Jager.

|| Sample fused on previously degassed young basalt.

¶ Accepted values have not been published.

* Typical weight 0.15 g.

Two samples of BKT-2u were collected independently, from the Bouroukie wadi, a tributary to the Kada Hadar drainage (Fig. 3). Each sample was gently wet-disaggregated to preserve the original grain sizes. The samples were washed for 10 min in dilute (~4%) HF, and sonified to rid the minerals of adhering clay particles, thoroughly rinsed with de-ionized H₂O and air dried. They were subjected to density, sieving, and magnetic separations.

BKT-2u contains a simple, euhedral, volcanic phenocryst assemblage consisting of 25% anorthoclase, 1% aegerine-augite, and common magnetite octahedra and zircon. The phenocrysts are encased in a matrix of collapsed pumice clasts. Small, 0.05–0.10 mm, magnetite octahedra and zircon crystals are common inclusions in the larger, 0.50–1.50 mm anorthoclase and aegerine-augite phenocrysts⁵. In addition, BKT-2u contains 0.5–2.0 mm, angular fragments of vesicular basaltic glass previously misidentified as obsidian, and basaltic rock fragments, perhaps incorporated during a mixed magma eruption¹. Anorthoclase and zircon were used for K/Ar and fission track dating, respectively^{1,4–6}.

One of the collected BKT-2u samples, sample 1, was subjected to a very refined size, density and magnetic separation to isolate anorthoclase of different physical properties that could be tested for K/Ar age concordance.

The first K/Ar age reported for BKT-2u was 2.63±0.05 Myr BP (ref. 5), which is the mean and 1σ standard deviation of four measurements of the same light, non-magnetic, 35/80 mesh anorthoclase fraction from sample 1 (tracers 277, 279, 297, and 311; Table 2). Additional measurements were reported on two finer size fractions of sample 1 anorthoclase (tracers 470 and 471; Table 2) and on a dense, magnetic separate in which each anorthoclase crystal contains a small (~0.5 mm) octahedral magnetite inclusion (tracer 472; Table 2)⁶. We also report two new K/Ar measurements made on the second BKT-2u sample, sample 2 (tracers 874 and 876; Table 2) and three measurements for a 35/80 mesh anorthoclase fraction separated from BKT-2L (tracers 466, 467 and 468; Table 2). Note that sample 2, BKT-2u, was not subjected to such a refined density and magnetic separation as sample 1.

In all, 14 conventional K/Ar measurements have been made on BKT-2. These are listed in Table 2, corrected for new decay, abundance and tracer constants. The 14 age determinations yield a mean age and standard deviation of the measurements of 2.93±0.11 Myr BP with an absolute range 2.75–3.12 Myr.

Note, however, that the K/Ar measurements of sample 2, BKT-2u, yield a slightly older set of ages, from 3.02 to 3.12 Myr. This compares with the range of sample 1 measurements from 2.75 to 2.98 Myr. The mean and standard deviation for sample 2 and sample 1 are 3.07±0.05 and 2.88±0.08 Myr BP, respectively. Zircon fission track data show that BKT-2u contains two zircon age populations; a dominant one around 2.7 Myr BP and a minor component around 5.0 Myr BP (ref. 1). This suggests that the older K/Ar measurements on sample 2 could be due to trace amounts of possible contaminating feldspars associated with the older zircon component. The younger set of K/Ar ages

measured for sample 1 could be due to the exclusion of detrital feldspars because of the more refined laboratory separation process that sample 1 underwent. However, even sample 1 feldspars may have suffered some contamination.

The concordance of all K/Ar measurements for various fractions of sample 1 argues that any contaminating effect is minimal. The concordant age group, 2.88±0.08 Myr BP, includes measurement of the magnetic anorthoclase fraction in which every feldspar grain contains an octahedral magnetite inclusion. We feel that this inclusion trait fingerprints the BKT-2 U and L eruptions, and is too rare for the tephra to contain contaminants of similar grains from an older eruption. Finally, we stress that the lever arm exerted on the measured age by detrital contamination is much less at Hadar than at most East African sites. This is because the sediments at Hadar are dominantly volcanoclastic⁹ with measured K/Ar and fission track ages on the detrital components being less than 30 Myr (ref. 6).

We accept the 2.88±0.08 Myr BP mean age as the most precise K/Ar measurement of the BKT-2 eruption. The fission track age of the dominant younger component of BKT-2u is 2.7±0.2 Myr BP (ref. 1), which is within the uncertainty of the K/Ar age but further work is needed to improve its precision.

Kadada Moumou basalt

The Kadada Moumou basalt (KMB) crops out along the Ounda Hadar wadi (Fig. 3), in the upper part of the Sidi Hakoma (SH) Member of the Hadar Formation (Fig. 1). The correlation of KMB with the fossil hominid localities in the lower SH Member, such as A.L. 200, is not precisely established. The SH Member

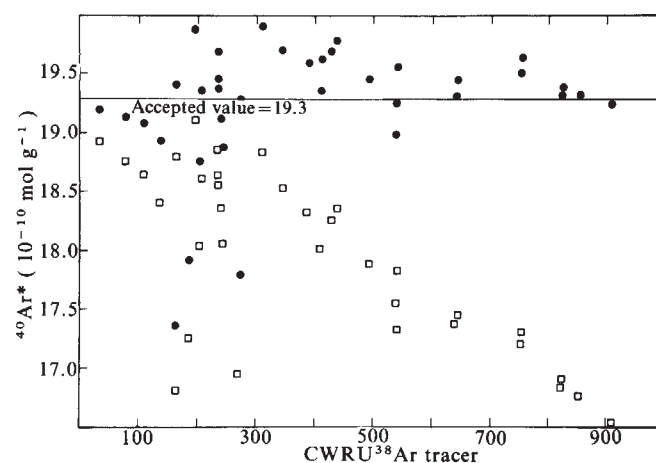


Fig. 2 Monitoring the depletion of the CWRU (Case Western Reserve University) ³⁸Ar tracer bulb against the LP6 biotite interlaboratory standard. Method A (□) utilizes glassblower-stipulated ³⁸Ar in the initial fill and an 18.6×10⁻¹⁰ mol g⁻¹ of ⁴⁰Ar* value for LP6. Method B (●) uses the machinist stipulation of the pipette volume and 19.3×10⁻¹⁰ mol g⁻¹ of ⁴⁰Ar* for LP6. Method A is wrong, method B is correct.

Table 2 Conventional K/Ar data for BKT-2 feldspars

Sample no., fraction, tracer run no.	Sample weight (g)	% K ₂ O	⁴⁰ Ar* ($\times 10^{-10}$ mol g ⁻¹)†	⁴⁰ Ar*/ ⁴⁰ Ar _T	Age (Myr)‡
BKT-2u, sample 1					
Sink 5, non-magnetic, 35/80 mesh					
277	3.8000	1.96	0.0778	0.77	2.75 ± 0.08
279	5.3850	2.00	0.0814	0.85	2.82 ± 0.08
297	4.1901	1.96	0.0810	0.80	2.86 ± 0.08
317	4.3815	1.96	0.0829	0.67	2.94 ± 0.08
873	2.0007	1.90	0.0793	0.40	2.89 ± 0.08
Sink 5, non-magnetic, 100/140 mesh					
470	4.0000	1.83	0.0789	0.73	2.98 ± 0.08
Sink 5, non-magnetic, 80/100 mesh					
471	4.0000	1.87	0.0795	0.68	2.94 ± 0.08
Sink 2, magnetic, 35/80 mesh§					
472	4.0000	1.73	0.0730	0.77	2.93 ± 0.08
BKT-2u, sample 2					
Sink 3, non-magnetic, 35/60 mesh					
874	2.0004	1.77	0.0769	0.34	3.02 ± 0.08
Sink 3, non-magnetic, -170 mesh					
876 Step 1 argon extraction	2.0001	1.85	0.0823	0.39	3.09 ± 0.08
Step 1 + higher temperature step	—	—	0.0823	0.39	3.09 ± 0.08
BKT-2L, Sink 4, non-magnetic, 35/80 mesh					
466	4.0000	1.47	0.0624	0.62	2.95 ± 0.08
467	4.0000	1.47	0.0585	0.68	2.76 ± 0.08
468	4.0000	1.47	0.0603	0.60	2.85 ± 0.08
					Mean s.d. = 2.92 ± 0.11

† Data corrected for new ³⁸Ar depletion constant.‡ Data corrected for new ⁴⁰K decay and abundance constants.

§ Each grain contains a magnetite octahedron.

thickens from ~70 m at Kada Hadar, where no basalt occurs but where numerous hominid localities exist, to ~160 m at Ounda Hadar and Ourda, where the basalt crops out, but where no hominids are found (Figs 1 and 3). KMB is probably stratigraphically above the lower SH Member hominid finds, as judged by (1) proportioning the stratigraphical thicknesses between the Kada Hadar and Ounda Hadar–Ourda sections; (2) by tentative correlation of gastropod-bearing layers above most of the hominid levels with similar layers below the basalt; and (3) by correlation of a prominent palaeomagnetic polarity transition, from normal to reverse, that occurs below the basalt in the Ounda Hadar section and above the gastropod layers in both the Ounda Hadar and Kada Hadar sections (Fig. 1)⁵. The KMB is, however, stratigraphically below such important hominid localities as A.L. 333 and A.L. 288 in the Denen Dora (DD) and Kada Hadar (KH) Members, respectively, because the KMB is definitely below the Triple Tuff (TT) and Ostracod clay marker horizons which form the base of the DD Member (Fig. 1).

Initial chronological measurements for the Hadar Formation centred around the KMB^{4,5}. This lava is a moderately low

potassium (~0.6% K₂O) tholeiite, which displays various stages of alteration depending on position in the 2–3 m thick flow. The base of the flow is conformable with sediments of the Hadar Formation. The sediments were baked to a depth of 2–3 cm, and the base of the flow was supercooled forming a 2–3 cm thick chilled zone. Protruding upwards from this contact are abundant bent 15-cm long calcite-filled pipe vesicles. These contact features suggest that the lava erupted over very shallow water or over water saturated sediments. The sedimentary environment over which the lava flowed is interpreted as a floodplain or marsh⁸. The reaction between lava and water-saturated sediments probably caused severe hydration of the lava¹.

Type A and B Kadada Moumou basalts

Recovery and dating of a new, unique basalt sample from the Kadada Moumou flow is a major factor which caused us to revise our minimum age estimate of 3.0 ± 0.2 (refs 4, 5) to 3.60 ± 0.15 Myr (ref. 1). The unique variety of this flow is termed 'type B'; all previously published data for KMB were accomplished on 'type A' varieties. These two varieties can be distinguished petrographically as well as geochronologically.

In general, the KMB is not ideal for dating because of variable alteration of the basaltic glass and mesostasis to a nontronite clay. The alteration presumably developed diagenetically after initial hydration of the flow, as described above. Being aware of the variable amount of clay alteration in KMB⁵, we conducted a detailed search of the flow and recovered a uniquely least-altered sample, KM76-5. This sample was collected *in situ* from the core of an oblate, spherical structure about 20 cm in diameter, within the middle of the 3-m thick flow. The core of the sample consists of dense, fine-grained basalt that grades radially outward to a zone containing minor flattened, nontronite-filled amygdulites. The rim of the structure contains abundant, flattened, and rounded nontronite-lined and filled vesicles. This structure conceivably originated as a cooled block of the lava flow reincorporated into the flow. Its major and trace element chemistry is identical to the rest of the flow (Table 3). The petrography of type B KM76-5 is, however, unique in that it is more nearly holocrystalline (Table 4) and essentially all of its glass is unaltered. In addition, the plagioclase and pyroxene of the type B variety are finer grained than type A varieties (Table 4) and the magnetite is equant (cubic) rather than rod-shaped

Table 3 Whole rock chemical data* for the Kadada Moumou basalt

Type	A-1 (n = 3)	A-2 (n = 4)	B (n = 3)
SiO ₂	51.16	52.01	52.13
Al ₂ O ₃	13.89	13.77	13.73
FeO	8.80	8.47	9.71
Fe ₂ O ₃	4.16	4.20	3.41
TiO ₂	2.92	2.94	2.90
MnO	0.23	0.25	0.24
MgO	4.23	3.97	4.36
CaO	7.87	8.47	7.74
Na ₂ O	3.67	3.58	3.57
K ₂ O	0.74	0.70	0.69
P ₂ O ₅	0.64	0.66	0.62
L.O.I.	0.75	0.92	0.49
Total	99.52	99.76	99.28
Rb	23	23	22
Sr	296	303	298
Ni	40	40	43
Zr	365	376	392
Ba	312	317	285
Rb/Sr	0.077	0.076	0.074
K/Rb	280	285	270

* X-ray fluorescence.

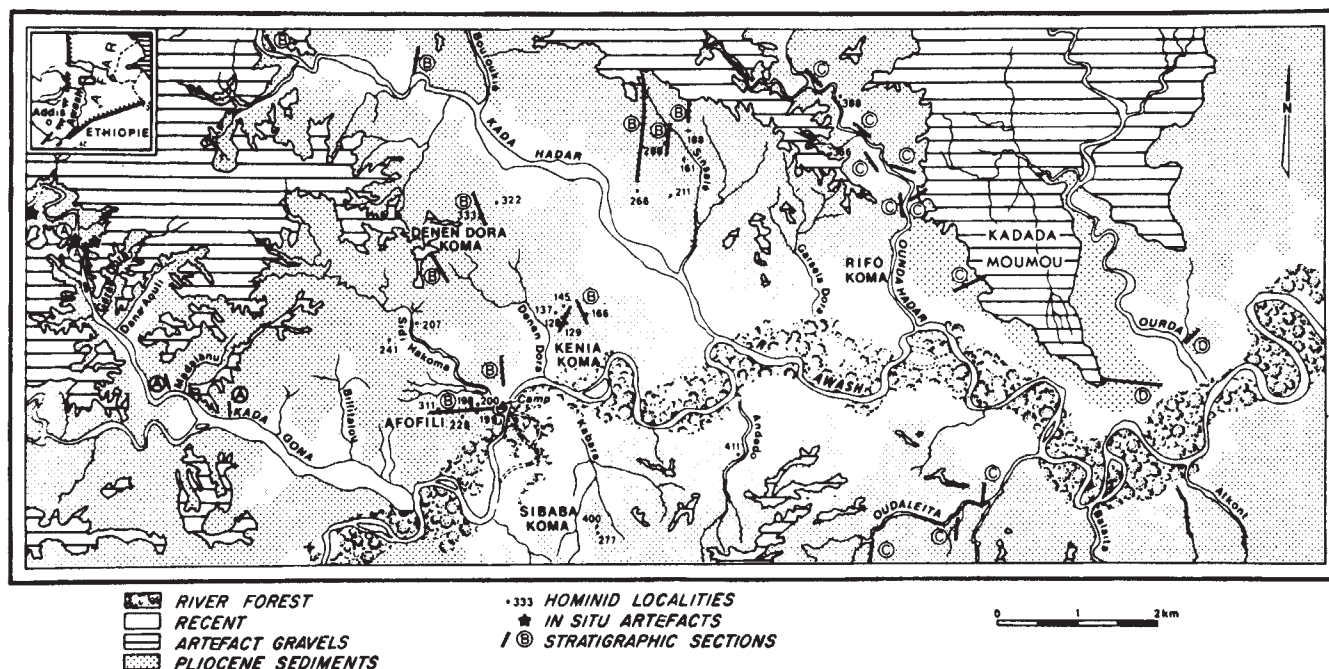


Fig. 3 Map of the Hadar hominid site showing location of stratigraphical sections and hominid localities (taken from ref. 8).

(acicular). Type A samples contain significantly greater amounts of glass, altered glass, and nontronite than type B samples, and their crystals are much coarser grained (Table 4).

Initially, a random piece of KM76-5 was prepared for K/Ar analysis. Subsequently, two additional samples were obtained by sectioning the core of the boulder into thin slabs. The sample was then sawn into two fractions, one labelled F (fair), which contained minor nontronite-filled vesicles, and the other labelled G (good), which was less vesicular. The nontronite in the type B basalt is thought to be a secondary amygdale fill and not an alteration product of the basaltic glass because, in thin section, the nontronite can be seen to be relegated to vesicles, and virtually none of the matrix is altered.

K/Ar data

The resultant mean and standard deviation of eight measurements of the KM76-5F sub-sample is 3.64 ± 0.12 Myr BP. For five measurements of the KM76-5G sub-sample this is 3.54 ± 0.12 Myr BP. Three runs of the original bulk sample, KM76-5, yield a mean value of 3.67 ± 0.10 Myr BP. Thus, the 16 measurements of the three sub-samples of the type B variety are internally concordant with a mean age of 3.61 ± 0.13 Myr BP.

In addition, type B samples produced uniformly higher proportions of radiogenic argon, ~25%, compared with only

~15% radiogenic argon for type A samples (Table 5). This enhancement of radiogenic argon is mainly due to a reduction in contaminating atmospheric argon, the additional radiogenic argon being a minor component. Typically, we and others observe the general correlation that the per cent radiogenic argon yield is inversely proportional to the amount of secondary alteration^{1,10-12}. We interpret the lack of alteration of the type B sample as due to the dense core structure preventing access of altering solutions.

In contrast to the type B variety of the flow, the type A basalts produce uniformly lower measured ages (Table 5, revised for new constants and tracer factor) with some ages as young as 3.0 Myr (revised; petrographical type A-1) and others higher at 3.3 Myr (revised; type A-2). K₂O and Ar measurements on physical separates of the sugary-textured² type A-1 basalt were attempted to see if a single sample could be shown to be internally discordant between low density glass-rich, nontronite-rich fractions and heavy glass-poor, nontronite-poor fractions. Initially, a rough separate into two density fractions was performed on type A-1 sample KM-2-74 (lab) and the heavier fraction was analysed (sample 1, Table 6). A large separate piece (sample 2) was then disk pulverized and subjected to a careful size/density separation into eight fractions (sample 2, Table 6). Modal counts were made on grain mount

Table 4 Modal mineralogy of type A and type B Kadada Moumou basalts and type A density separates

KM2-74 (lab) (type A-1)											
Grain mounts											
Type and no. analysed	A-1 (n = 3)	A-2 (n = 3)	B (n = 3)	Sink 1 100/120	Sink 2 100/120	Sink 3 100/120	Sink 4 100/120	Sink 1 80/100	Sink 2 80/100	Sink 3 80/100	Sink 4 80/100
Mode*											
Plagioclase*	14	10	8	8	17	39	20	13	27	28	24
Clinopyroxene*	13	3	2	50	20	3	0	46	12	10	2
Opaque*	6	4	1	6	8	6	2	9	6	5	3
Plagioclase†	13	18	29	6	9	16	12	6	11	9	8
Pyroxene†	12	11	33	7	6	11	3	3	15	12	4
Opaque†	5	7	15	4	3	5	3	3	4	t	4
Glass†	17	32	7	16	31	11	45	17	20	30	41
Alteration†‡	15	10	5	2	7	9	14	3	5	6	14
Vesicles†	2	1	1								
Calcite†	t	3	0								

Mode is based on 1,000 counts. * Phenocryst phases (>0.05 mm). † Groundmass phases (<0.05 mm); t, trace. ‡ Includes vug and intersertal fill.

Table 5 Conventional K/Ar data for the Kadada Moumou basalt

Sample and tracer no.	Type	Sample weight (g)	% K ₂ O	⁴⁰ Ar* ($\times 10^{-12}$ mol g ⁻¹)	⁴⁰ Ar*/ ⁴⁰ Ar _T	Age (Myr)‡
KM-2-74 (lab)						
258	A-1	10.0001	0.764	3.405	0.14	3.09 ± 0.25
860	A-1	7.9988		3.345	0.14	3.08 ± 0.26
KM-3-74						
257	A-1	10.0001	0.747	3.110	0.26	2.88 ± 0.20
262†	A-1	10.0000		3.210	0.15	2.98 ± 0.25
263	A-1	10.0000		4.112	0.24	3.80 ± 0.22
865	A-1	8.0041		3.252	0.18	3.02 ± 0.20
KM75-123-3						
339	A-1	7.9992	0.673	2.633	0.16	2.71 ± 0.25
360	A-1	7.3755		2.826	0.19	2.86 ± 0.25
840	A-1	7.0004		2.910	0.15	2.92 ± 0.25
KM-1-73						
261	A-2	9.9997	0.657	3.007	0.17	3.21 ± 0.23
KM-1-74 (lab)						
259	A-2	10.0001	0.727	3.538	0.15	3.38 ± 0.24
859	A-2	10.0000		3.534	0.15	3.38 ± 0.24
KM4-74						
256	A-2	10.0001	0.686	3.237	0.14	3.28 ± 0.23
260	A-2	10.0001		3.269	0.13	3.38 ± 0.26
866	A-2	8.0039		3.292	0.12	3.36 ± 0.28
KM76-5						
609	B	7.0017	0.656	3.484	0.21	3.68 ± 0.25
683	B	7.0041		3.576	0.21	3.78 ± 0.25
787	B	7.0051		3.358	0.18	3.55 ± 0.25
KM-76-5G						
658	B	7.0009	0.657	3.165	0.19	3.34 ± 0.24
786	B	7.0027		3.349	0.19	3.54 ± 0.24
812	B	7.0063		3.507	0.24	3.70 ± 0.22
867	B	10.0001		3.359	0.21	3.56 ± 0.22
868	B	10.0003		3.363	0.21	3.56 ± 0.22
KM76-5F						
659	B	7.0093	0.659	3.416	0.25	3.59 ± 0.22
716	B	7.0035		3.452	0.27	3.78 ± 0.22
813	B	6.9999		3.460	0.27	3.65 ± 0.21
835	B	9.9939		3.459	0.29	3.65 ± 0.20
841	B	7.0045		3.418	0.27	3.61 ± 0.22
861	B	9.9999		3.187	0.31	3.36 ± 0.20
869	B	10.0051		3.540	0.23	3.72 ± 0.22
870	B	10.0003		3.609	0.25	3.80 ± 0.21

† Not baked out.

‡ New constants and tracer depletion factor.

Table 6 Size/density fractionation experiment KM2-74 (lab) type A-1

Sample and fraction	³⁸ Ar tracer no.	Weight (g)	K ₂ O %	⁴⁰ Ar* ($\times 10^{-12}$ mol g ⁻¹)	⁴⁰ Ar*/ ⁴⁰ Ar _T	Age (Myr)
Sample 1						
12/35 mesh, raw, run 1	258	10.0	0.76	3.40	0.153	3.09 ± 0.20
12/35 mesh, raw, run 2	860	8.0	0.76	3.34	0.140	3.08 ± 0.20
80/100 mesh gross heavy fraction	899	8.1	0.73	3.85	0.146	3.63 ± 0.28
Sample 2						
80/100 mesh						
Raw, unfractionated	1,078†	8.5	0.787	3.81	0.130	3.36 ± 0.27
Sink 1 (heaviest)	1,073†	9.0	0.525	2.53	0.116	3.33 ± 0.31
Sink 1, repeat run	1,092	10.1		2.69	0.141	3.55 ± 0.32
Sink 2	1,074†	9.0	0.971	4.33	0.131	3.09 ± 0.26
Sink 3 + 4 combined	1,075†	11.0	1.139	4.95	0.135	3.02 ± 0.28
100/120 mesh						
Sink 1	1,076†	5.5	0.413	2.13	0.101	3.57 ± 0.41
Sink 2 (no Ar run)	—	—	0.807	—	—	—
Sink 3	1,077†	7.0	1.073	4.82	0.138	3.13 ± 0.28
Sink 4 (no Ar run)	—	—	1.122	—	—	—

† Unbaked.

thin sections (Table 4) and 6 fractions were analysed for K₂O and Ar (Table 6).

The heavier fractions, being poor in glass and nontronite and rich in pyroxene (Table 4), are considerably lower in K₂O (0.4–0.5% K₂O). The light fractions are rich in glass, nontronite, and plagioclase (Table 4) and high in K₂O (1.07–1.12% K₂O; Table 6). The heaviest fractions (sink 1) produce uniformly older ages, with four measurements having an average and standard deviation of 3.5 ± 0.5 Myr. For this separation experiment most of the separates were not pre-baked and yielded a low per-

centage of radiogenic argon that resulted in a large internal imprecision in the individual argon measurements.

These results, though imprecise, indicate that the altered type A basalt is internally discordant. The data associate younger ages with more abundant glass, a phase that is prone to argon loss, or with more nontronite that is, in part, observed to be an alteration product of glass in type A basalts.

The possibility that the type B Kadada Moumou basalt has some excess magmatic ⁴⁰Ar has to be considered. We think this is unlikely, however, because of the rarity of examples of excess

argon in aphanitic terrestrial basalts¹⁵. The flow is extensive and thin enough that it should have thoroughly degassed. Furthermore, there is no evidence for older acidic crust in western Afar¹ that could have contaminated the magmas as was indicated for the New Zealand basalts studied by McDougall *et al.*¹³; no xenoliths of cognate inclusions have ever been found in either the Kadada Moumou flow or numerous other western Afar flows¹. It is also unlikely that the flow picked up boulders or fragments of older basalt because the flow erupted onto over-bank sediments of clay and clay-sized particles. On the contrary, there is a distinct inverse correlation of whole rock measured ages with degree of observed sample alteration and, in the type A fractionation experiment, with glass and nontronite content (Tables 4–6). Alteration typically causes ⁴⁰Ar loss (and K gain) in K/Ar systems. Therefore, we regard the spectrum of ages obtained from the Kadada Moumou flow as being due to alteration variably affecting the K/Ar age measurement. We regard the 3.60 ± 0.15 Myr BP age measurement of the type B basalt, the least altered variety of the flow, as a best-estimate minimum age for the flow. However, even the type B age requires some qualification because of the small, but perhaps significant, amount of alteration also seen in this variety.

Accepting the older 3.60 Myr age for the Kadada Moumou basalt necessitates a revised correlation of the Hadar palaeomagnetic sequence to the palaeomagnetic time scale: the details will be considered elsewhere. The Kadada Moumou flow and

overlying and underlying sediment are reversed. The revised age places the reverse sequence within the Gilbert Epoch which ended 3.4 Myr BP instead of in the Mammouth Event 3.15–3.05 Myr BP as previously interpreted (Fig. 1).

Conclusion

Adoption of new decay and abundance constants, recalibration of the ³⁸Ar laboratory tracer pipette and data from several new samples requires us to revise the K/Ar ages for the BKT-2 tephra, from 2.65 to 2.9 Myr BP, and the Kadada Moumou basalt, from 3.0 to 3.6 Myr BP. All the fossil hominid remains at Hadar are found below BKT-2, and many hominid localities occur stratigraphically below the Kadada Moumou basalt. Therefore, on the basis of this revised chronology, the fossil hominid material in the Hadar Formation is interpreted to range from at least 2.9 Myr BP to somewhat older than 3.6 Myr BP. These revised ages permit a correlation of the pre-basalt Sidi Hakoma Member hominid fossils at Hadar with the footprints and fossils at Laetoli which is 700 km to the south and well dated between 3.6 and 3.75 Myr BP (new constants)¹⁶.

This work was supported by NSF. We thank S. A. Mertzman and Franklin and Marshall College for use of XRF facilities, and Derek York for reading the manuscript. Rick Cotman, Bob Sedivy and Larry Dasch helped with the K/Ar and mineral separation laboratory. We thank Karen Toil for helping to prepare this manuscript.

Received 13 July; accepted 29 December 1981.

1. Walter, R. C. thesis, Case Western Reserve Univ. (1980).
2. Steiger, R. H. & Jäger, E. *Earth planet. Sci. Lett.* **36**, 359–362 (1977).
3. Mankinen, E. A. & Dalrymple, G. B. *J. geophys. Res.* **84**, 615–626 (1979).
4. Taieb, M., Johanson, D. C., Coppens, Y. & Aronson, J. L. *Nature* **260**, 289–293 (1976).
5. Aronson, J. L. *et al. Nature* **267**, 323–327 (1977).
6. Aronson, J. L., Walter, R. C., Taieb, M. & Naeser, L. W. *Proc. VIII Pan-African Cong., Nairobi* (eds Leakey, R. E. & Ogot, B. A.) (National Museums of Kenya, 1980).
7. Alexander, E. C., Jr., Mickelson, G. M. & Lanphere, M. A. *Short Pap. 4th int. Conf. on Geochronology* (ed. Zartman, R. E.) (Open File Rep. 78–701, 1978).
8. Taieb, M. & Tiercelin, J. J. *Bull. Soc. Geol. Fr.* **1**, 243–253 (1979).
9. Aronson, J. L. & Taieb, M. in *Hominid Sites: Their Paleoenvironmental Settings* (eds Rapp, G. & Vondra, C.) (American Association for the Advancement of Science, 1981).
10. Mankinen, E. A. & Dalrymple, G. B. *Earth planet. Sci. Lett.* **17**, 89–94 (1972).
11. Seidemann, D. *Geochim. cosmochim. Acta*, **42**, 1721–1734 (1978).
12. Dalrymple, G. B., Lanphere, M. A. & Clague, D. A. *Init. Rep. DSDP* **55**, 659–676 (1981).
13. McDougall, I., Polach, H. A. & Stipp, J. J. *Geochim. cosmochim. Acta* **33**, 1485–1520 (1969).
14. Baksi, A. K. *Earth planet. Sci. Lett.* **21**, 431–438 (1974).
15. Dalrymple, G. B. *Earth planet. Sci. Lett.* **6**, 47–55 (1969).
16. Leakey, M. D. *et al. Nature* **278**, 317–323 (1976).

Initiation of ice ages by creep instability and surging of the East Antarctic ice sheet

Gerald Schubert* & David A. Yuen†

* Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

† Department of Geology, Arizona State University, Tempe, Arizona 85287, USA

Creep or shear heating instability is shown to be a plausible means of initiating a surge of the East Antarctic ice sheet and plunging the Earth into an ice age. The occurrence of this instability requires only that the ice thickness exceed a critical value which may not be much larger than the present ice dome thickness. Creep instability can occur on the fast time scale of 100–1,000 yr required by inferences of past rapid cooling events and sudden rises of sea level before the onset of Northern Hemisphere glaciations.

THE Antarctic surge theory¹ for the inception of ice ages invokes an instability in the ice sheet to melt its base and cause it to slide into the ocean, thereby forming a huge ice shelf around the southern continent. This would increase the Earth's albedo and lead to global climatic cooling and the inception of glaciation in the Northern Hemisphere. It would also produce a rapid rise in sea level by as much as 15–20 m at the end of an interglacial². Several lines of evidence indicate that these events did indeed occur near the end of the last interglacial^{3–6}; in particular, the association of the >8 m sea level rise ~120,000 yr ago inferred from raised coral reefs in New Guinea with the 2 °C cooling of the surrounding ocean on a 1,000 yr time scale deduced from oxygen isotope studies of giant clam shells strongly supports the surge theory⁶. The magnitude of the sea level rise requires the surge of a large fraction of the

East Antarctic ice sheet⁷; disintegration of the West Antarctic ice sheet would increase sea level by only ~4 m (refs 8, 9).

The only physical mechanism which seems capable of initiating a surge of the East Antarctic ice dome, or major portions thereof, is creep instability due to frictional heating at the base of a deforming ice sheet^{10–16}. Such an instability could cause massive melting at the base of the ice sheet on a very rapid time scale and allow the ice to slide freely over the underlying bedrock. The possibility of melting by shear heating is a consequence of the strong temperature *T* dependence of the rheology of ice. Frictional heating in the basal shear layer of the ice sheet tends to raise the basal temperature and reduce the viscosity near the base. If the thickness of the ice sheet remains constant, so will the basal shear stress, and the increase in temperature and decrease in viscosity will tend to increase

the rate of frictional heating. Enhanced heating will further increase the temperature in a potentially unstable manner. Shear heating instability can occur if the ice sheet thickness h exceeds a critical value h_c in which case steady, subsolidus sliding of the ice sheet is not possible¹⁵.

The scenario we develop calls on climatically enhanced accumulation to increase the thickness of the East Antarctic ice sheet above the critical value necessary for the onset of explosive creep instability. The climatic change could be a consequence of latitude-dependent variations in solar insolation that accompany changes in the Earth's orbital motion. Thus our model may provide a link between the astronomical^{17,18} and Antarctic surge theories for the inception of ice ages by using the former to initiate the latter. However, if the Antarctic surge phenomenon does have a crucial role then it is not necessary for orbital forcing to cause major spreading of ice sheets. Any climatic occurrence resulting in increased accumulation over Antarctica and a sufficiently large thickening of the Antarctic ice sheet might initiate surging and the onset of an ice age.

Model

The critical thickness h_c required for creep instability and surging can be calculated using a one-dimensional, steady-state, thermomechanical model for the gravitational creep of a constant thickness h ice sheet down a slope making an angle α to the horizontal¹⁵. The coordinate perpendicular to the slope is y ; the ice surface is $y = 0$ and the base is $y = h$. The surface of the ice and its base have the same slope in this model, while the surface and bedrock slopes of the East Antarctic ice sheet are unequal. This difference between the mathematical model and the real ice sheet is unimportant, however, as the shear stress τ in an ice sheet is mainly determined by the surface slope¹⁹. The distribution of shear stress in our model ice sheet,

$$\tau = -\rho g y \sin \alpha \quad (1)$$

where ρ is the density of ice and g is the acceleration due to gravity, is thus representative of the shear stress in the real ice sheet as long as we equate α with the local surface slope.

The equations governing the temperature T in the model ice sheet and its downslope creep velocity u are¹⁵

$$\frac{d^2 T}{dy^2} - \frac{v}{\kappa} \frac{dT}{dy} + \frac{2A}{k} \{\rho g y \sin \alpha\}^4 \exp\left(\frac{-E^*}{RT}\right) = 0 \quad (2)$$

$$\frac{du}{dy} = -2A \{\rho g y \sin \alpha\}^3 \exp\left(\frac{-E^*}{RT}\right) \quad (3)$$

where k is the thermal conductivity and κ is the thermal diffusivity. The temperature equation (2) is a balance between conduction, vertical advection with velocity v , and viscous dissipation. Equation (3) is the creep law of ice²⁰, A is a rheological parameter, E^* is the activation energy and R is the universal gas constant. The vertical velocity is given by the accumulation rate. We use two models for the y -dependence of v : $v = v_c$ = constant and $v = v_c(1 - y/h)$. The constant v model is preferred because v decreases from v_c to 0 in a thin boundary layer at the base of the ice sheet¹⁵. Equations (2) and (3) are solved subject to the conditions

$$T = T_0, \quad \text{on } y = 0 \quad (4)$$

$$k \frac{dT}{dy} = q_b, \quad u = 0, \quad \text{on } y = h \quad (5)$$

where T_0 is the constant temperature of the ice surface and q_b is the geothermal heat flux entering the base of the ice sheet. We have assumed that there is no basal sliding. This two-point boundary value problem is solved using standard numerical procedures¹⁵.

Parameter values

The East Antarctic ice sheet has surface slopes of 0.05–0.10° and thicknesses of 2–4 km (refs 21–24); accumulation rates are typically 0.03 m yr⁻¹ in the interior and an order of magnitude larger near the coast^{14,25}. Experimental data on the creep of polycrystalline ice suggest that E^* is in the range 60–70 kJ mol⁻¹ at temperatures ≤ -10 °C (refs 26–28). The value of A depends on grain size, impurity content and crystal orientation. We have used $A = 8.75 \times 10^{-13} \text{ s}^{-1} \text{ Pa}^{-3}$, a value representative of the experimental determinations^{27,28}. We have also used $k = 2.5 \text{ W m}^{-1} \text{ K}^{-1}$, $\kappa = 1.3 \text{ mm}^2 \text{ s}^{-1}$, $\rho = 900 \text{ kg m}^{-3}$, $g = 9.8 \text{ m s}^{-2}$, $T_0 = 223 \text{ K}$, and $q_b = 40 \text{ mW m}^{-2}$. The value of T_0 is representative of the annually averaged surface temperature in the interior of East Antarctica and the value of q_b is a typical continental heat flux²⁹.

Critical thickness for shear heating instability

Nonlinear coupling between velocity and temperature results in the solution curves shown in Fig. 1 for $E^* = 60 \text{ kJ mol}^{-1}$. For given α and v_c there is a maximum thickness h_c for which steady creep is possible. The dashed portions of the curves indicate solutions which have basal temperatures at or above the melting temperature of ice. The maximum thickness increases with increasing accumulation rate (Fig. 1a) because downward advection of cold ice to the basal shear zone enhances the

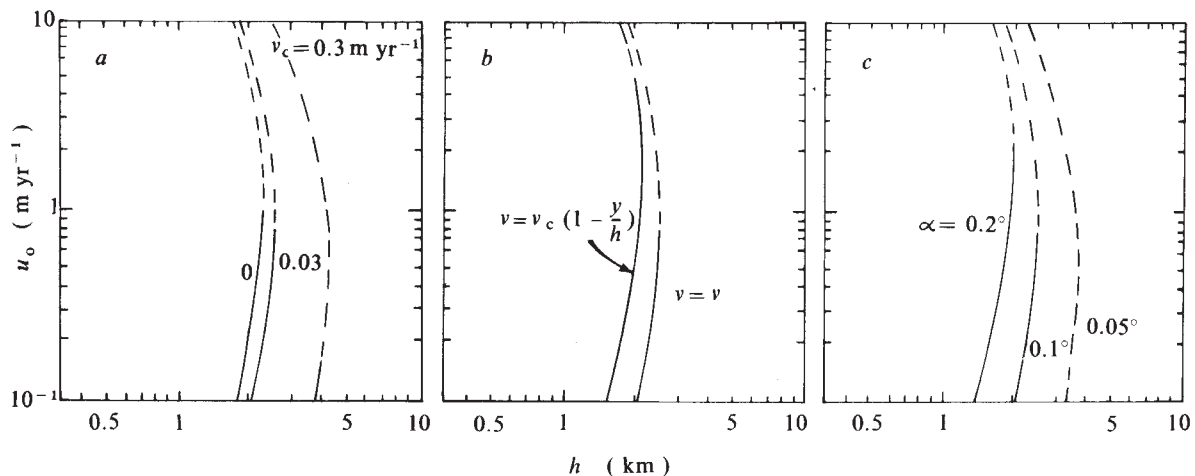


Fig. 1 Surface velocity u_0 as a function of ice sheet thickness for $E^* = 60 \text{ kJ mol}^{-1}$. *a*, Effect of varying the accumulation rate, $\alpha = 0.1^\circ$, $v = v_c$; *b*, influence of vertical velocity profile, $\alpha = 0.1^\circ$, $v_c = 0.03 \text{ m yr}^{-1}$; *c*, effect of varying surface slope, $v_c = 0.03 \text{ m yr}^{-1}$, $v = v_c$. For given α and v_c there is a maximum thickness for which stable, steady, gravitational sliding is possible.

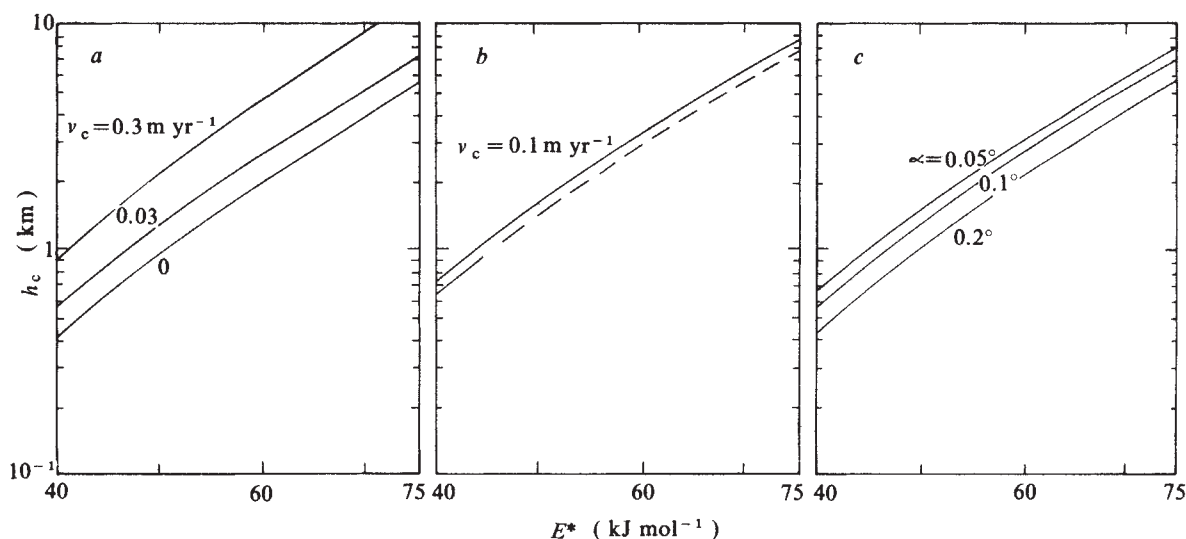


Fig. 2 Maximum thickness h_c for stable, steady, downhill creep of an ice sheet as a function of activation energy E^* . *a*, Effect of varying the accumulation rate, $\alpha = 0.1^\circ$, $v = v_c$; *b*, influence of vertical velocity profile, $\alpha = 0.1^\circ$, $v = v_c$ (solid), $v = v_c(1 - y/h)$; *c*, effect of varying surface slope, $v_c = 0.03 \text{ m yr}^{-1}$, $v = v_c$.

importance of frictional heating. For the same reason, h_c in the constant v model is larger than it is in the situation where v decreases linearly with depth (Fig. 1*b*). A decrease in α also leads to an increase in h_c (Fig. 1*c*) because large ice thicknesses are then necessary to produce the shear stresses required for internal deformation. Basal temperatures for the supercritical states are generally at or above the melting point of ice. Even the subcritical states have basal melting when the ice thickness is large. The tendency of downward advection to lower basal temperatures is more than compensated by the associated increase in viscous dissipation.

Figure 2 shows the dependence of the maximum thickness on activation energy; h_c increases approximately exponentially with E^* for given α and v_c . The thicknesses within the East Antarctic ice sheet (2–4 km) are near the calculated values of h_c for $\alpha = 0.1^\circ$, $v_c = 0.03 \text{ m yr}^{-1}$ and E^* between 60 and 70 kJ mol^{-1} .

Occurrence of instability and surging

The existence of a maximum thickness for which steady, stable, subsolidus, gravitational sliding of an ice sheet is possible is the fundamental reason for the occurrence of catastrophic shear heating instability. According to Fig. 2, for example, an ice sheet on a 0.1° slope would creep down hill in a steady, stable manner if its thickness were $\leq 5 \text{ km}$ for an activation energy of 70 kJ mol^{-1} and an accumulation rate of 0.03 m yr^{-1} . However, if its thickness were suddenly increased above 5 km by a rapid climatically enhanced accumulation it could no longer slide stably. An instability would occur in which frictional heating near the base of the ice would result in massive melting and a surge of the ice sheet, as shown in Fig. 3. The maximum thickness of the East Antarctic ice sheet is $\sim 4 \text{ km}$, its activation energy may be $\sim 70 \text{ kJ mol}^{-1}$, surface slopes of about 0.1° occur over widespread areas, and typical accumulation rates are 0.03 m yr^{-1} . Thus it is possible that the East Antarctic ice sheet lies close to the point of creep instability.

The occurrence of shear heating instability and surging in response to a rapid increase in ice thickness $h > h_c$ is a consequence of the super-exponential time scale on which thermal runaways of this nature take place³⁰. The growth times for such explosive finite-amplitude instabilities far exceed those calculated on the basis of linear theory^{31,32}. The very rapid growth of the instability limits the time available for the ice sheet to self adjust to the increase in accumulation in ways that might avoid instability. For example, if the flux could be maintained constant by thinning the ice sheet, it would be stabilized, because the solutions for flows subjected to a constant flux

boundary condition are single-valued in contrast to the multi-valued solutions occurring for the constant thickness boundary condition^{33,34}. However, the sudden enhancement in accumulation is externally imposed, the ice sheet must thicken initially, and the catastrophic occurrence of shear heating instability precludes self adjustment and the maintenance of constant flux^{35,36}. Simple energy arguments support the reality of the instability³⁷.

The fast time scale associated with finite-amplitude shear heating instability is essential to ice sheet surging on a continental scale; the isotopic evidence⁶ for rapid surging at the end of the last interglacial has already been cited. It is probably not feasible for basal melting in subcritical conditions to lead to large-scale disintegration of the Antarctic ice sheet, as originally suggested by Wilson¹, because there is no fast time scale involved. The ice sheet has sufficient time to flow and accommodate itself to changing surface conditions so as to preclude massive collapse by large-scale decoupling between the ice and the bedrock. This is supported by inferences from modelling^{38,39} and the discoveries of large sub-ice lakes²³ that large areas of the base of the East Antarctic ice sheet are at the melting temperature. Despite this, the ice sheet is apparently stable. This does not preclude localized basal sliding or surging due to melting by shear heating. Basal sliding at the rate of 10 m yr^{-1} has been reported for the ice sheet in the Mizuko Plateau⁴⁰,

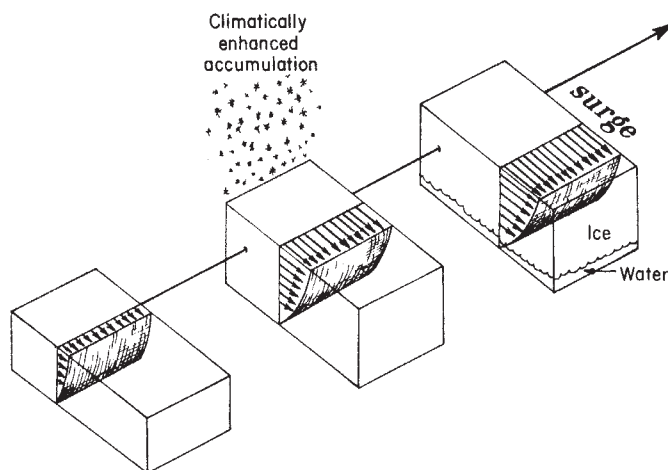


Fig. 3 Sketch of the events leading to shear heating instability and surging of the East Antarctic ice sheet.

for example. Such localized sliding contributes to the global stabilization of the present ice sheet.

The above modelling results suggest that past ice ages may have been caused by shear heating instability of the East Antarctic ice sheet. A climatically enhanced accumulation is required to thicken the ice beyond its maximum thickness for stable, steady, gravitational sliding. Any climatic event which enhances accumulation over Antarctica would suffice to initiate the process; variations in the Earth's orbital motion or enhanced volcanic activity are possibilities. The thickening of the ice must occur quickly so that flow within the ice cannot lead to an

Received 30 November 1981; accepted 20 January 1982.

1. Wilson, A. T. *Nature* **201**, 147–149 (1964).
2. Hollin, J. T. *Nature* **208**, 12–16 (1965).
3. Hollin, J. T. *Quat. Res.* **2**, 401–408 (1972).
4. Hollin, J. T. *Boreas* **6**, 33–52 (1977).
5. Hollin, J. T. *Nature* **283**, 629–633 (1980).
6. Aharon, P., Chappell, J. & Compston, W. *Nature* **283**, 649–651 (1980).
7. Denton, G. H., Armstrong, R. L. & Stuiver, M. in *The Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 267–306 (Yale University Press, 1971).
8. Mercer, J. H. *Int. Ass. Sci. Hydrol. Publ.* **79**, 217–225 (1968).
9. Hughes, T. J. *J. geophys. Res.* **78**, 7884–7910 (1973).
10. Robin, G. de Q. *J. Glaciol.* **2**, 523–532 (1955).
11. Robin, G. de Q. *Can. J. Earth Sci.* **6**, 919–928 (1969).
12. Budd, W. *ANARE Sci. Rep. Publ.* **108**, 216 pp. (1969).
13. Bozhinskiy, A. N. *Izv. Akad. Nauk. SSSR Mekh. Zhidk. Gaza* **5**, 169–172 (1975).
14. Clarke, G. K. C., Nitsan, U. & Paterson, W. S. B. *Rev. Geophys. Space Phys.* **15**, 235–247 (1977).
15. Yuen, D. A. & Schubert, G. *J. Glaciol.* **24**, 195–212 (1979).
16. Cary, P. W., Clarke, G. K. C. & Peltier, W. R. *Can. J. Earth Sci.* **16**, 182–188 (1979).
17. Milankovitch, M., in *Handbuch der Klimatologie* (eds Koeppen, W. & Geiger, R.), 1–176 (Gebrüder Borntraeger, Berlin, 1930).
18. Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
19. Nye, J. F. *J. Glaciol.* **2**, 103–107 (1952).
20. Glen, J. W. *Proc. R. Soc. A* **228**, 519–538 (1955).

accommodation and stabilization of the ice sheet against the changing conditions. Instability and large-scale melting must occur extremely fast for the same reasons. Because shear heating instability is explosive in nature, such fast collapse of the ice sheet should be possible. Theoretical calculations of the super-exponential growth rates accompanying these catastrophic instabilities yield the necessary fast time scales.

D.A.Y. thanks the NATO Postdoctoral Fellowship Program, NSERC of Canada, and the Research Corporation for financial support. This research was also supported by NSF grant DPP80-23723.

21. Heezen, B. C., Tharp, M. & Bentley, C. R. *Antarctic Map Folio Series—Folio 16* (American Geographical Society, New York, 1972).
22. Drewry, D. J. *Polar Rec.* **17**, 359–374 (1975).
23. Robin, G. de Q., Drewry, D. J. & Meldrum, D. T. *Phil. Trans. R. Soc.* **B279**, 185–196 (1977).
24. Ravich, M. *Nauka Zhizn*, No. 1, 112–115 (1976).
25. Paterson, W. S. B. *The Physics of Glaciers* 2nd edn (Pergamon, Oxford, 1981).
26. Mellor, M. & Testa, R. *J. Glaciol.* **8**, 131–145 (1969).
27. Weertman, J., in *Physics and Chemistry of Ice* (eds Whalley, E., Jones, S. J. & Gold, L. W.) 320–337 (Royal Society of Canada, Ottawa, 1973).
28. Shoji, H. & Higashi, A. *J. Glaciol.* **24**, 487–489 (1979).
29. Roy, R. F., Blackwell, D. D. & Decker, E. R. in *The Nature of the Solid Earth* (ed. Robertson, E. C.) 506–543 (McGraw-Hill, New York, 1972).
30. Kassoy, D. R. *Quat. J. Mech. appl. Math.* **30**, 71–89 (1977).
31. Yuen, D. A. & Schubert, G. *EOS* **60**, 391 (1979).
32. Yuen, D. A. & Peltier, W. R. in *Physics of the Earth's Interior* (eds Dziewonski, A. & Boschi, E.) 432–463 (North-Holland, Amsterdam, 1980).
33. Fowler, A. C. *J. Glaciol.* **25**, 183–184 (1980).
34. Fowler, A. C. & Larson, D. A. *Geophys. J. R. astr. Soc.* **63**, 333–345 (1980).
35. Schubert, G. & Yuen, D. A. *J. Glaciol.* **25**, 355 (1980).
36. Clarke, G. K. C. & Paterson, W. S. B. *J. Glaciol.* **25**, 355–356 (1980).
37. Weertman, J. & Birchfield, G. E. (in preparation).
38. Sugden, D. *Arctic alp. Res.* **9**, 21–47 (1977).
39. Hughes, T. in *The Last Great Ice Sheets* (eds Denton, G. H. & Hughes, T.) 222–275 (Wiley, New York, 1981).
40. Mae, S. J. *J. Glaciol.* **24**, 53–61 (1979).

LETTERS TO NATURE

Is the Sun an oblique magnetic rotator?

G. R. Isaak

Department of Physics, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK

The recent observation¹ of rotational splitting of the non-radial p modes of $l=1$ and $l=2$ (with $n \sim 20$) of the 5-min global oscillation of the Sun was interpreted in terms of rotational splitting associated with a rapid rotation of the solar interior. The precise value deduced for the angular velocity Ω of the interior depends on the assumed variation of Ω with depth and on the weighting function. Thus with a weighting function² of the form

$$\bar{\Omega} = \int \frac{\Omega(r)}{V(r)} dr \int \frac{dr}{V(r)} \quad (1)$$

where $V(r)$ is the local velocity of sound at radius r , one obtains values of $\Omega_{\text{core}}/\Omega_{\text{surface}}$ ranging from 2 to 9 as the radius at which Ω , assumed to be constant for smaller r , ranges from 0.6 to 0.15 of the solar radius. Such a weighting function proportional to the time the wave spends at any particular radius seems very plausible for the effect of any parameter, be it angular velocity or magnetic field, on the propagation of nearly plane waves as they bounce between the centre and the surface. Here I attempt to correlate the size of the observed rotational splitting with the enigmatic 12.2-day variation in the measurement of solar oblateness discovered by Dicke³ previously⁴ and to interpret the observed near equality of the amplitudes of the m -components of the $l=1$ and $l=2$ rotationally split modes. This leads to the first empirical evidence that an asymmetric magnetic rotator with megagauss magnetic fields exists in the interior of the Sun. It also suggests that the magnetic energy of young stars is a sizeable fraction of their gravitational energy.

I speculate, together with Dicke³, that this observed periodicity and the rotational period of the core of the Sun are identical.

This assumption allows one to refine the model of the solar interior by demanding that up to a radius r the Sun rotates uniformly, say, at the sidereal rate of $0.95 \mu\text{Hz}$ (corresponding to 12.2-days) and beyond that radius at the surface rate of $0.46 \mu\text{Hz}$ and that the integral (1) equals the observed rotational splitting of about $0.78 \mu\text{Hz}$ (sidereal). On the above assumptions the integral can be satisfied for a current solar model⁵ with the boundary at $r=0.9 R_{\odot}$. It is natural to speculate that this value corresponding to the boundary between the rapidly rotating core and the slowly rotating surface layers does measure the depth of the convection zone. A slower transition from the rapidly rotating core to the more slowly rotating surface layers would imply a deeper convection zone.

The observation¹ of $l=1$ triplets with nearly equal components (and less conclusively $l=2$ quintuplets with nearly equal components) does not accord with the expectation of purely rotational splitting of solar modes being viewed from Earth nearly along the solar equator. Thus, excluding the possibility of frequent excitation from randomly distributed sources and a correspondingly large damping, the expected skew symmetry of the central $m=0$ component about the solar equator would lead one to expect that the $m=0$ component ought to be absent in a global view of the Sun and very weak for a tilt of the solar equator of some 6° with respect to that condition as obtains during the observing period.

The presence of this component strongly suggests a perturbation which makes a large angle with the axis of rotation and rotates the angular distribution pattern through a substantial angle. Although purely hydrodynamic effects⁶ at the boundary between the rapidly rotating core and the slower outer layer could be responsible the required velocities would have to be high ($\sim 1 \text{ km s}^{-1}$) and in substantially different directions from those of the rotating matter. I, therefore, suggest that the perturbation is magnetic and in the absence of a theory of the combined action of rotation and magnetic field on frequencies of oscillations, that the analogy between rotational and magnetic splittings in an acoustic spectrum and magnetic and quadrupole splittings of atomic and nuclear energy levels be explored.

Such considerations suggest that an intense internal magnetic field, such as might be associated with an oblique (asymmetric?) magnetic rotator, could be responsible for the perturbation. Assuming a linear superposition of rotational and magnetic effects leads one to expect^{7,8} frequency shifts $\Delta\nu_{n,l,m}$ superimposed on acoustic frequencies ν of the order of

$$\left(\frac{\Delta\nu}{\nu}\right)_{n,l,m} = \alpha_{n,l,m} \frac{E_M}{|E_G|} \quad (2)$$

where E_M is the magnetic energy and E_G the gravitational energy of the Sun and $\alpha_{n,l,m}$ is of the order of unity and depends on the precise configuration of the magnetic fields and on the mode (n, l, m) . The positions of the unshifted lines cannot be measured nor, at present, calculated. Upper limits on the magnetic energy can, nevertheless, be estimated through equation (2) from the symmetry of the observed multiplets together with the assumption that the obliqueness angle with respect to the rotational axis is near 90° (ref. 3). The $l = 1$ triplets, observed between $n = 16$ and $n = 26$, are symmetric to within the measured linewidth of $0.6 \mu\text{Hz}$ that is

$$\delta(\Delta\nu_{n,l}) = \langle \nu_{n,l,0} - \frac{1}{2}(\nu_{n,l,1} + \nu_{n,l,-1}) \rangle \leq 0.6 \times 10^{-6} \text{ Hz} \quad (3)$$

This implies that for $|E_G| = 6 \times 10^{48} \text{ erg}$ and $E_M = (B^2/8\pi)\bar{V}$ where B is the magnetic field and $\bar{V}$ is the volume of the Sun, excluding the convection zone, $\langle B^2 \rangle_{\text{max}}^{1/2} = 3.3 \times 10^6 \text{ G}$, if

$$\bar{\Delta\alpha} = \langle \alpha_{n,l,0} - \frac{1}{2}(\alpha_{n,l,1} + \alpha_{n,l,-1}) \rangle = 1 \quad (4)$$

whereas if $\bar{\Delta\alpha}$ is substantially below unity the upper limit will be correspondingly higher.

A lower limit is given by demanding that the angular distribution rotates during the observation time¹ $\tau = 28$ days by at least 1 rad that is $2\pi\Delta\nu\tau > 1$, yielding a lower limit $\langle B^2 \rangle_{\text{min}}^{1/2} = 1.1 \times 10^6 \text{ G}$, if $\alpha = 1$.

If the earlier identification concerning the significance of the 12.2-day variation in the measurement of solar oblateness is assumed the value τ ought to be reduced to 12 days and $2\pi\Delta\nu\tau > 1$ yields $\langle B^2 \rangle_{\text{min}}^{1/2} = 1.7 \times 10^6 \text{ G}$, if $\alpha = 1$.

The observable effects of the magnetic field are, of course, weighted over the solar interior and this would be expressed in the size of the factor $\alpha_{n,l,m}$. In the absence of such model calculations one might put $\alpha_{n,l,m} = 1$ as a crude approximation and weight the square of the magnetic field in the solar interior $B^2(r)$ with a weighting function of the form of equation (1). Rough calculations show that the resultant is an observable mean square magnetic field which is smaller or larger than the volume averaged $B^2(r)$ depending on whether the field is extended or decreases rapidly (for example, as the matter density) with radius.

Speculations concerning large internal magnetic fields have been made before⁹⁻¹¹ but this is the first clear empirical evidence pertaining to the interior of the Sun which suggests their presence.

Such intense fields presumably have a fossil origin⁹. Although such fields are but a small fraction of the maximum fields allowed by the extended virial theorem¹², the much higher angular velocities of young solar type stars¹³ suggest that this may not have been the case in the early evolutionary stages of the Sun and presumably other main sequence stars. The magnetic flux, and therefore the internal field, is likely to be decreasing with time because of ohmic dissipation⁹ and also because of flux loss and angular momentum loss through the surface layers through the solar wind^{10,14,15}. In the early life of the Sun the magnetic energy may have been a non-negligible fraction of the total energy of the star.

Although the magnetic fields in the solar core considered above are large their contribution to the pressure in the core is likely to be too small to be significant in the context of the solar neutrino problem¹⁶.

The mixing that such magnetic fields are likely to provide¹⁰ will, however, probably also solve the problems of the missing solar neutrinos^{17,18} as well as the Earth's climatic problems caused by an ancient cool Sun¹⁹.

The asymmetric oblique magnetic rotator has a nutation frequency¹⁰ ω given by

$$\frac{\omega}{\Omega} = \frac{E_M}{|E_G|} = \frac{1}{\Delta\alpha} \frac{\delta(\Delta\nu_{nl})}{\nu_{nl}} \leq \frac{1}{\Delta\alpha} \frac{0.6 \mu\text{Hz}}{3 \text{ mHz}} \leq \frac{1}{\Delta\alpha} 2 \times 10^{-4}$$

where Ω is the angular velocity of the core taken to be $6 \times 10^{-6} \text{ s}^{-1}$. Thus the period of one nutation should be longer than $166 \Delta\alpha \text{ yr}$ and as $\Delta\alpha$ is likely to be < 1 it is tempting to speculate, together with Dicke²⁰, that this corresponds to the 22-yr solar cycle. With this identification, the magnetic energy of the Sun would be increased above the value given by equation (4) and $\langle B^2 \rangle^{1/2}$ would be $9 \times 10^6 \text{ G}$.

Other aspects of solar activity which seem to be explained by the above model will be discussed elsewhere.

If the above ideas concerning internal magnetic fields correspond to reality the Sun will fit quite naturally into the sequence of young Ap stars to white dwarfs and pulsars. Moreover the Sun will be able to sustain a very much richer spectrum of modes of vibration and it is conceivable that the 160-min oscillation^{21,22} is such a magnetohydrodynamic mode.

The magnetic effects should give rise to asymmetries which will be detectable with the higher resolution of longer strings of velocity measurements in the near future and may also produce periodic displacements with 12.2 or 6.1-day periods.

If the Sun has substantial magnetic fields in the core it is tempting to speculate that other stars also do and it is suggested that the β Cephei stars, for example, show large magnetic rather than rotational splittings and the magnetic fields may destabilize the stars sufficiently²³ to make them pulsate.

I thank Professor H. H. Voigt for the hospitality and the Akademie der Wissenschaften for support during my stay in Göttingen.

Received 1 September; accepted 7 December 1981.

- Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. & Roca Cortes, T. *Nature* **293**, 443-445 (1981).
- Christensen-Dalsgaard, J. & Gough, D. O. Preprint (Univ. Cambridge, 1981).
- Dicke, R. H. *Sol. Phys.* **47**, 475-515 (1976).
- Dicke, R. H. & Goldenberg, H. M. *Phys. Rev. Lett.* **18**, 313-316 (1967).
- Allen, C. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1973).
- Fricke, K. J. & Kippenhahn, R. A. *Rev. Astr. Astrophys.* **10**, 45 (1972).
- Goossens, M. *Astrophys. Space Sci.* **43**, 9-18 (1976).
- Goossens, M. *Astrophys. Space Sci.* **44**, 397-404 (1976).
- Cowling, T. G. *Mon. Not. R. astr. Soc.* **105**, 166-174 (1945).
- Mestel, L. & Takhar, H. S. *Mon. Not. R. astr. Soc.* **156**, 419-436 (1972).
- Dicke, R. H. *Astrophys. J.* **228**, 898-902 (1979).
- Chandrasekhar, S. & Fermi, E. *Astrophys. J.* **118**, 116-141 (1953).
- Dicke, R. H. *Astrophys. J.* **171**, 331-362 (1972).
- Schatzman, E. *Ann. Astrophys.* **25**, 18-29 (1962).
- Dicke, R. H. *Nature* **202**, 432-435 (1964).
- Mestel, L. *Mon. Not. R. astr. Soc.* **140**, 177-196 (1968).
- Ezer, D. & Cameron, A. G. W. *Astrophys. Lett.* **1**, 177 (1968).
- Bahcall, J. N., Bahcall, N. A. & Ulrich, R. K. *Astrophys. Lett.* **2**, 91 (1968).
- Roxburgh, I. W. *Plains Feux Sur La Physique Solaire*, 21-23 (CNRS, Paris, 1978).
- Dicke, R. H. *New Scientist* **83**, 12-14 (1979).
- Brookes, J. R., Isaak, G. R. & van der Raay, H. B. *Nature* **259**, 92-95 (1976).
- Severny, A. B., Kotov, V. A. & Tsap, T. T. *Nature* **259**, 87-89 (1976).
- Chandrasekhar, S. & Limber, D. N. *Astrophys. J.* **119**, 10-13 (1954).

Inferring solar UV variability from the atmospheric tide

Neill S. Cooper

Atmospheric Physics Group, Department of Physics, Imperial College, London SW7 2BZ, UK

A new method of determining variations in the solar UV output is presented here. The method is based on the fact that the semi-diurnal (12 h) atmospheric tide is dominated by the component forced by ozone absorption of UV. This tide shows very little interannual variability; hence a variation in UV of 20%, as suggested by Heath and Thekaekara¹, would have an easily discernable effect on the tide. It can be shown that a change of 20% in UV would result in a 12% change in the semi-diurnal tide. The failure to find any solar cycle-related variation in the semi-diurnal tide suggests that the UV varies by <2% over the sunspot cycle.

The suggestion¹ that the solar UV output varies by ~20% over the 11 yr sunspot cycle is controversial^{2,3}. Such changes in UV would alter ozone concentrations, temperature and wind in the stratosphere^{4,5}. This may affect the vertical propagation of planetary waves in winter^{6,7} and hence the troposphere. One indicator of ozone layer heating whose variation over the sunspot cycle is on record, but which has received little attention, is the semi-diurnal (12 h) atmospheric tide. Unlike oceanic tides, this is due to the heating effect of the Sun⁸. The dominant component is from ozone absorption of UV⁸.

The model of atmospheric tides used here is that of Chapman and Lindzen⁸. The altitudinal and latitudinal dependences are separated. The latitudinal variation of the tides are Hough functions and the vertical structure equation, with appropriate boundary conditions, determines the magnitude and phase of each component of the tide at all levels. The effect of zonal winds and the interaction between different tidal components are ignored, which results in an underestimate of the amplitudes^{9,10}. The only forcings considered were the heatings due to ozone and water vapour absorbing solar radiation. The heating profiles were obtained using a model similar to that of Forbes and Garrett¹¹.

Data were obtained from two sources of 'hourly averaged annual pressure', that is, the mean pressure at each hour of the day averaged over 1 yr. Results for Macau were from Resultados das Observações Meteorológicas, 1952–78, while the Indian values were from Indian Weather Reviews, Annual

Table 1 Details of the tropical stations

Station	Location	Years used	Total (yr)
Bombay	18.9° N 72.8° E	1948–54, 1956–62	14
Calcutta	22.5° N 88.3° E	1948–54, 1956–63	15
Macau	22.2° N 113.6° E	1952–1975	24
Madras	13.1° N 80.3° E	1951–54, 1956–61	10
Nagpur	28.6° N 79.1° E	1949–54, 1956, 1948–63	13
New Delhi	18.5° N 77.2° E	1949–53, 1956–63	13
Poona	4.6° S 73.9° E	1948–54, 1956–62	14
Seychelles	55.5° E	1951–54	4

Summaries, 1948–63. Table 1 lists the eight stations used, their locations and the years for which their data were available.

The amplitudes of the mean tides, averaged over the 107 station years, are 0.809 mbar (diurnal) and 1.180 mbar (semi-diurnal). These are larger than the values the theory predicts, as a result of ozone and water vapour forcing alone, by factors of 5 and 1.6 respectively. This result is consistent with that of previous work in atmospheric tidal modelling^{8,10}. The causes of the disagreement between theory and observation are the approximation of linearity in the theory^{9,10}, and the neglect of local tropospheric heat sources^{12,13}. The latter is particularly evident in the diurnal tide, which is much larger over the continents than over the oceans.

Having obtained results for mean conditions, the heating profile is then altered by changing UV and ozone concentrations to examine the effect of UV variability. The ozone concentration is altered, due to changes in UV, by the amount predicted by stratospheric models^{4,5}. This has a negligible effect on the tides, as there is complete absorption of UV by ozone. The UV variation over the sunspot cycle is parameterized by a $\pm 10\%$ variation in the ozone heating in the Hartley Band, Huggins Band and Hertzberg continuum. The model predicts a change of $\pm 6.2\%$ in the semi-diurnal tide in the tropics, while the diurnal (24 h) tide would vary by $\pm 0.8\%$. It is assumed that the component of the tides unexplained by the theory has the same fractional change with UV variation as the theoretical component. As the semi-diurnal tide shows very little annual variation, a change of 12% would be readily visible.

Table 2 The mean tide, and the predicted and observed increase in the amplitude from sunspot minimum to sunspot maximum

	Diurnal (μ bar)	Semi-diurnal (μ bar)
Mean	809	1,180
Predicted	14	148
Observed	-41 ± 13	5 ± 11

Results are from the stations listed in Table 1. The error estimates are one standard error.

The results for all the stations are considered together to reduce the standard errors. They are normalized at each station, before averaging over all the stations, so that changes solely in amplitude add together linearly and not vectorially.

The results are shown in Table 2 which highlights the discrepancies between the theoretical predictions and the observations. The diurnal tide shows a much larger variation ($-41 \pm 13 \mu$ bar) over the sunspot cycle than predicted (14 μ bar). The semi-diurnal tide, on the other hand, is predicted to increase by 48 μ bar but is observed to increase by only $5 \pm 11 \mu$ bar. This latter result agrees with previous work¹⁴, where no correlation was found between sunspot activity and the amplitude of the semi-diurnal tide at Batavia (for 1866–1945) and at Manila (1890–1938).

The variation in the diurnal tide is actually larger than the whole component forced by ozone heating, thus it is unlikely to be due to variations in UV. Hence, between 1948 and 1975 there was a diurnally varying tropospheric parameter which had a significant change over the sunspot cycle. However, where a significant change was expected, in the semi-diurnal tide, none was found. The observations of this tide provide an estimate of the maximum value for the sunspot cycle variation in UV. As shown above, a 20% change in the UV would cause a 12.3% change in the semi-diurnal tide, so assuming linearity, the observed tidal variation is due to a $0.5\% \pm 1.6\%$ variation in the UV. This result is more precise than any derived using conventional techniques³. Analysis of a larger surface pressure data set may lead to even greater precision.

This result shows that the atmospheric semi-diurnal tide can be used to examine possible solar UV variations. Indeed, this tide may also be usable for remote sensing of the stratosphere, in a way analogous to those for examining the internal structure of the Earth¹⁵ and Sun¹⁶. If this is the case, the large exant quantity of hourly surface pressure data might be used to determine possible stratospheric variations (unrelated to solar activity) over the past 100 yr.

I thank other members of the Atmospheric Physics Group at Imperial College for their support, particularly Dr J. S. A. Green, also the Meteorological Office Library for providing data. The work was financed by the NERC and the Appleton Laboratory.

Received 23 September 1981; accepted 13 January 1982.

1. Heath, D. F. & Thekaekara, M. P. *The Solar Output and its Variations* (ed. White, O. R.) (Colorado Associated University Press, 1977).
2. Mount, Rottman & Timothy *Geophys. Res. Lett.* (in the press).
3. Brasseur, G. & Simon, P. C. *J. geophys. Res.* (in the press).
4. Callis, L. B. & Nealy, J. E. *Geophys. Res. Lett.* **5**, 249–252 (1978).
5. Penner, J. E. & Chang, J. S. *Geophys. Res. Lett.* **5**, 817–820 (1978).
6. Bates, J. R. O. *J. R. met. Soc.* **103**, 397–430 (1977).
7. Geller, M. A. & Alpert, J. C. *J. atmos. Sci.* **37**, 1197–1215 (1980).
8. Chapman, S. & Lindzen, R. S. *Atmospheric Tides* (Reidel, Dordrecht, 1970).
9. Lindzen, R. S. & Hong, S.-S. *J. Atmos. Sci.* **31**, 1421–1446 (1974).
10. Walterscheid, R. L., De Vore, J. G. & Venkataswaran, S. V. *J. atmos. Sci.* **37**, 455–470 (1980).
11. Forbes, J. M. & Garrett, H. B. *Geophys. Res. Lett.* **5**, 1013–1016 (1978).
12. Lindzen, R. S. *Mon. Weath. Rev.* **106**, 526–533 (1978).
13. Hamilton, K. *Mon. Weath. Rev.* **109**, 3–17 (1981).
14. Haurwitz, B. et al. *J. geophys. Res.* **62**, 489–491 (1957).
15. Backus, G. E. & Gilbert, J. F. *Phil. Trans. R. Soc. A* **266**, 123–192 (1970).
16. Christenson-Dalsgaard, J. & Gough, D. O. *Nature* **259**, 89–92 (1976).

Diurnal variation of mesospheric ozone

G. Vaughan

Meteorological Office, Bracknell, Berkshire RG12 2SZ, UK

Photochemical models¹⁻³ predict a considerable variation in ozone concentration between day and night in the mesosphere. To investigate this, four Petrel rockets were flown from South Uist (57°22'N, 7°22'W) on 2 October 1979. Ozone was measured by observing the atmospheric attenuation of a narrow band of UV radiation, using interference filters to define a bandwidth of ~10 nm. The rockets were launched (1) at moonset, 02.00h, (2) at sunrise, 06.00h, (3) at 09.30h and (4) at 15.30h; round 1 observing moonlight and the other three observing sunlight. The first two used the occultation technique⁴ and contained two photometers with wavebands centred around 265 and 290 nm (Fig. 1), giving ozone information from 48 to 95 km. For rounds 3 and 4 the Sun was at a zenith angle of 70°, and a single waveband centred around 265 nm gave a height range of 44 to 64 km. Full details of the experiment will be published elsewhere. The first three rounds were successful, but the fourth did not give good data because the rocket nosecone failed to clear the field of view. The results show significant diurnal variation above 54 km, which exceeds a factor of 2 above 65 km and reaches a factor of 10 between night time and sunrise at 90 km.

To retrieve ozone density from the raw flight data, a profile was produced of attenuation as a function of height for each sensor. The signals near apogee showed no drift in detector sensitivity, but rounds 2 and 3 showed some dependence on rocket attitude (coning modulation). The observations near apogee were used to derive corrections for this, and to deduce the unattenuated signals for the rest of the flights. The attenuation profiles were completed by using a radar track of the first 64 s of flight, extrapolated using a ballistic trajectory, to give rocket height.

For zenith angle 70°, the transmission $\xi(h)$ at height h may be written

$$\xi(h) = \frac{\int d\lambda I_0(\lambda) S(\lambda) \exp\{-\sigma_{\text{oz}}(\lambda) N_{\text{oz}}(h) + \sigma_{\text{R}}(\lambda) N_{\text{air}}(h)\}}{\int d\lambda I_0(\lambda) S(\lambda)} \quad (1)$$

where I_0 is the unattenuated source irradiance at wavelength λ , $S(\lambda)$ the sensor sensitivity, $\sigma_{\text{oz}}(\lambda)$ the ozone absorption coefficient and $N_{\text{oz}}(h)$ the path total of ozone molecules between rocket and source. σ_{R} and N_{air} refer to Rayleigh scattering by air molecules. This equation may be solved for each value of ξ to obtain $N_{\text{oz}}(h)$, and the number density $n_{\text{oz}}(h)$ (given by

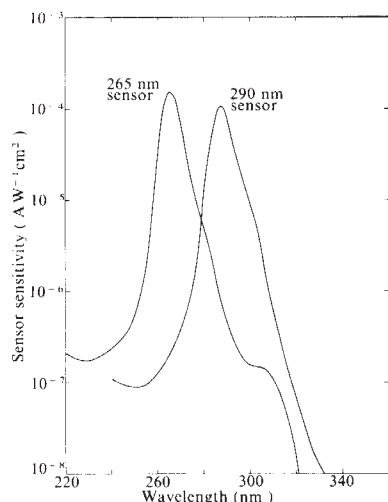


Fig. 1 Photometer sensitivities as a function of wavelength for round 2.

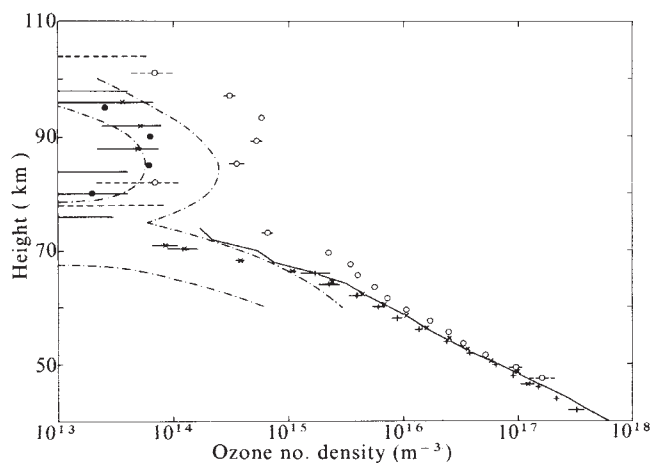


Fig. 2 Ozone number densities as a function of height. Error bars are random standard errors for the present measurements, and χ is the zenith angle of the centre of the Sun or Moon, viewed from the rocket. $\circ$, 02.00h, $\chi = 94.5$ (Moon); $\times$, 06.00h, $\chi = 95.5$ (Sun); $+$, 09.30h, $\chi = 69.7$ (Sun); $-$, daytime from ref. 14; $\bullet$, sunset from ref. 4; $- - -$, envelope of night-time values from ref. 16.

$-\cos \theta dN_{\text{oz}}/dh$ for zenith angle θ) may then be obtained by differentiating equation (1) with respect to h . The raw attenuation values for round 3 were smoothed by fitting a quadratic curve over 4-km intervals (about 30 individual observations), which also gave $d\xi/dh$ with its statistical standard error.

For rounds 1 and 2, the attenuation profile was first expressed as a function of tangent ray height. Because of the finite angular size of the Sun and Moon the attenuation at the rocket was a smoothed version of that from a point source. $N_{\text{oz}}(h)$ was retrieved at discrete levels assuming constant scale height between levels, by an iterative process similar to that described by Miller and Ryder⁴ and Miller⁵. Standard errors in $N_{\text{oz}}(h)$ were obtained by calculating the covariance matrix.

To deduce number densities from the path totals, ozone was assumed constant in a series of concentric layers, thus the path totals $N(h)$ could be written $\sum L_i(h) n_i$ where $L_i(h)$ is the length of the path through the i th layer and n_i the number density there. The inversion of this problem is very sensitive to noise in $N(h)$ and the method described by Rodgers⁶ was used to stabilize the solution without introducing appreciable bias. This involved using an 'a priori constraint' composed of a hand-smoothed version of an unconstrained retrieval. Below 70 km, ozone number densities were retrieved for 1 km layers but to stabilize the solution above 70 km, layers 4-km thick had to be used. In each case standard errors were obtained by calculating the covariance matrix.

For each sensor, the solar spectrum used was that of Broadfoot⁷ below 300 nm and Simon⁸ for wavelengths above 300 nm. Ozone absorption coefficients were taken from Inn and Tanaka⁹, corrected for temperature after Vigroux¹⁰. Rayleigh scattering was calculated using Penndorf's¹¹ value for cross-section, and the atmospheric profile provided by the Stratospheric and Mesospheric Sounder (SAMS) on the Nimbus 7 satellite (J. J. Barnett, personal communication). Solar limb darkening was calculated from data given by Moe and Milone¹².

Round 3 produced only one ozone profile, on the ascent leg, but rounds 1 and 2 also gave profiles on the descent leg. However, because of the need to extrapolate the radar data beyond 64 s, there is considerable uncertainty in the rocket position on the descent, and data from the ascents only are presented, with a height uncertainty of ~250 m. (These profiles were in good agreement with those of the descents.) The results are presented as number densities in Fig. 2 and mixing ratios in Fig. 3, atmospheric density being calculated from SAMS temperatures, extended above 75 km with the CIRA (1972)¹³ standard atmosphere for 60°N. Values shown are from the 290 nm sensors below 64 km and from the 265 nm sensors

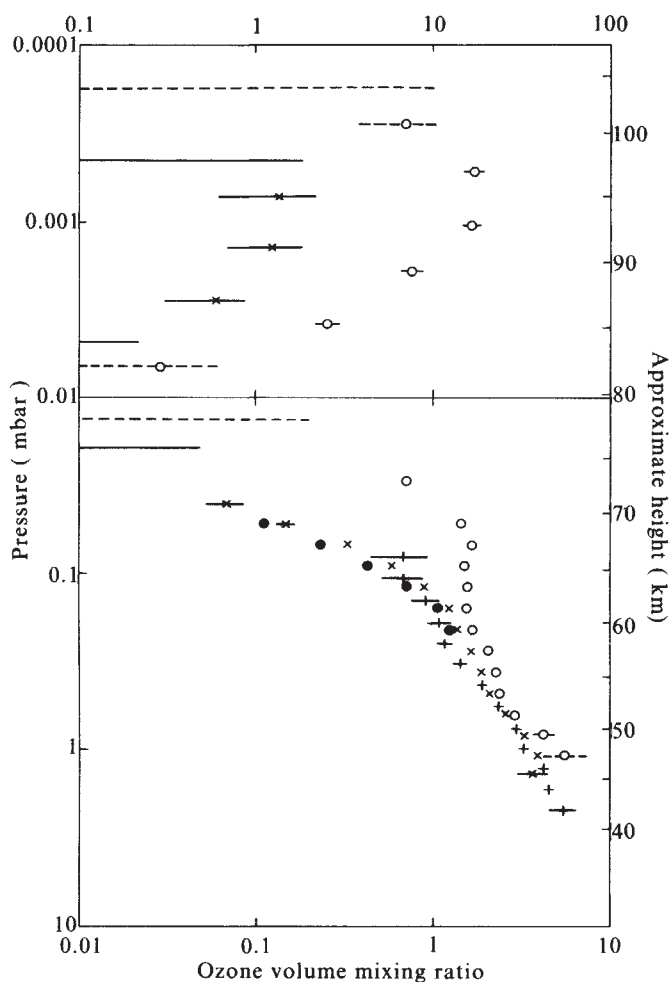


Fig. 3 Ozone mixing ratios as a function of pressure. Error bars are random standard errors. Note the scale change on the abscissa above 0.01 mbar. $\circ$, 02.00h; $\times$, 06.00h, assuming spherical symmetry in retrieval; $\bullet$, 06.00h, allowing for ozone variation along the path; $+$, 09.30h.

above this height. For rounds 1 and 2 the height resolution is ~ 5 km, and for round 3, ~ 4 km.

Errors in the ozone absorption coefficient (especially the temperature correction) and the spectral calibration contribute most of the systematic uncertainties in the results; the former amount to $\sim 4\%$ and the latter to 2% increasing to 5% below 48 km. Below 70 km for the dawn round, a test was made of the sensitivity of the retrieval to the hypothesis of horizontal uniformity, as ozone is predicted to vary rapidly near the terminator above 60 km. The photochemical model described below was used to predict the ratio of the concentration in each layer to its value at 90° zenith angle, and these ratios were incorporated in the retrieval. The differences from the original profile are shown in Fig. 3—they increase from $\sim 5\%$ at 58 km to a maximum of 28% at 67 km.

The results show a significant day–night difference in ozone concentration above 53 km, which reaches a factor of 10 above 70 km. The retrieval technique was unable to give more than an upper limit on the ozone amounts in the ‘dip’ of the profiles around 78 km, but the large differences persist in the upper ozone layer around 90 km. The dawn and morning profiles agree with Krueger and Minzner’s¹⁴ (45° N, daytime) standard midlatitude ozone model between 46 and 66 km, and the dawn values above 80 km agree well with those of Miller and Ryder⁴ (58° N, September, sunset) and Llewellyn and Witt¹⁵ (68° N, March, sunset, not shown on Fig. 2). The night-time measurements are compared above 60 km with the envelope of equatorial profiles obtained using stellar occultation by Hays

and Roble¹⁶. Unlike the present values, these show a divergence from the Krueger and Minzner¹⁴ model below 65 km, which is not supported by other workers^{17,18}. Above 65 km, our values are the greater at all altitudes, with a peak concentration 9 km higher, possibly due to the difference in latitude.

Previous measurements of the day–night difference of ozone below 65 km have been reported by Hilsenrath¹⁹ (38° N, March) and Anderson *et al.*²⁰ (50° S, December), both of whom show a factor of 2 at 65 km, in agreement with the present work. The change in column total above 50 km of 10^{16} molecules cm^{-2} agrees with the measurement of Penfield²¹, even though the profiles he showed were very different from ours. Above 70 km, our night-time column total of $1.24 \pm 0.07 \times 10^{15} \text{ cm}^{-2}$ agrees well with that of Wilson and Schwartz²², who report $1.1 \pm 0.4 \times 10^{15} \text{ cm}^{-2}$. These authors, like Penfield, used ground-based microwave spectroscopy, which has the disadvantage of poor height resolution (~ 20 km below 70 km), precluding a direct comparison with our measurements at specific heights. However, they show no sign of the minimum in ozone concentration around dawn and small maximum in mid-morning above 65 km as predicted by theory (see below). In contrast, the present measurements show the dawn profile crossing that at 09.30 near 65 km, indicative of just such a feature.

Above 80 km, there have been no coincident measurements of day and night ozone profiles to compare with this work. Model predictions of day–night differences in this region vary considerably, and depend on the value adopted for the vertical diffusion coefficient¹. Hunt²³ predicts a night-time enhancement factor of 2 and Shimazaki and Laird² predict 100. Theoretical predictions below 70 km are less variable but a quantitative comparison between theory and experiment requires a model run at the same latitude with the same profiles of relevant atmospheric variables. Comparison will therefore be made with

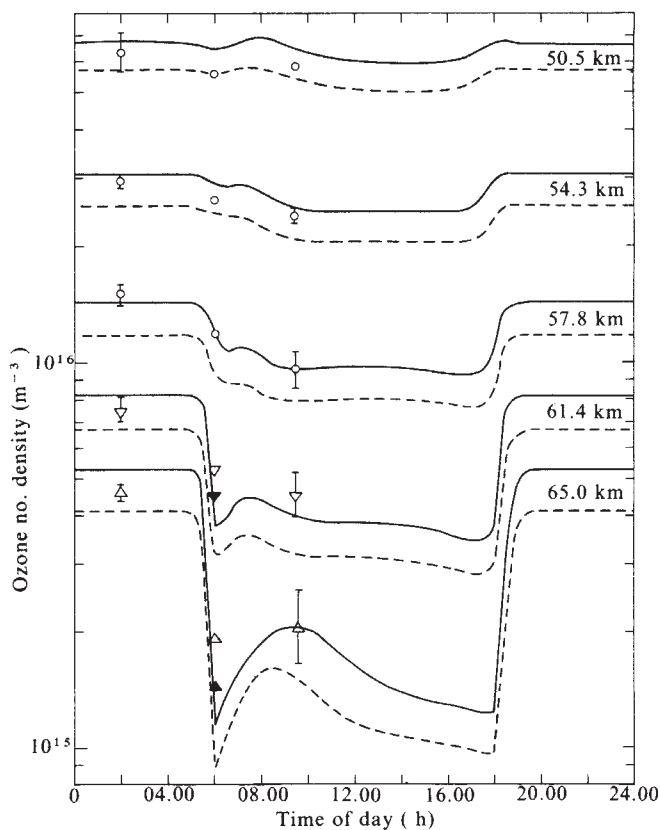


Fig. 4 Comparison between observations and theory. Solid lines are theoretical calculations with 2 p.p.m. water vapour and the broken lines correspond to 4 p.p.m. water vapour. Open symbols at dawn refer to retrievals assuming symmetry and closed symbols refer to retrievals with variable ozone (see text and Fig. 3).

a one-dimensional model adapted from the Oxford University two-dimensional model²⁴. This integrates the photochemical continuity equations diurnally (with no transport processes) for the odd oxygen, hydrogen, nitrogen and chlorine active species and their associated reservoirs, using rate constants as given in JPL 81.3 (ref. 25). A basic timestep of 5 min (less around dusk and dawn) is used for levels half a scale height (~ 4 km) apart in the vertical. For this work, the solar spectrum and absorption coefficients were based on the recommendations of Nicolet^{26,27} with the parameterization of Frederick and Hudson²⁸ being used for water vapour photodissociation. Temperatures were taken from SAMS data, the latitude fixed at 57° N and the time of year at equinox. A full description of the modelling work will be published elsewhere.

Two sets of model calculations, corresponding to water vapour mixing ratios of 2 and 4 p.p.m., are compared with the observations in Fig. 4. The basic day-night difference in ozone above 55 km is caused by the recombination of atomic oxygen to form ozone at dusk ($O + O_2 + M \rightarrow O_3 + M$) and its formation by ozone photodissociation at dawn. The small peak in mid-morning above 60 km results from the depletion of water dissociation products (H, OH, HO_2) during the night and their slow release from water during the morning; these products are the main destruction agents for odd oxygen in the mesosphere.

The agreement between theory and observation is consistent with a water vapour mixing ratio profile falling from ~ 3 p.p.m. at 50 km to ~ 2 p.p.m. at 65 km. The magnitude of the diurnal variation (which is relatively insensitive to the water vapour) is overestimated at the upper levels by $\sim 15\%$, possibly due to uncertainties in the rate constants used—the quoted uncertainty²⁵ for the rate of $O + O_2 + M \rightarrow O_3 + M$, for instance, is enough to account for this discrepancy. It may be concluded, then, that the observations are in agreement with current photochemical theory (given the uncertainty in the latter) provided that mesospheric water vapour abundances of 2–3 p.p.m. are accepted. Recent measurements of this quantity have varied from 1.5 p.p.m. (ref. 29) to 15 p.p.m. (ref. 30) at 60 km, and this work supports a mixing ratio near to the lower end of this range.

I thank J. H. Seymour for work on the design and preparation of the experiments, also S. G. Palmer and F. H. Price for contributions to the data analysis, and D. E. Miller for support and advice.

Received 25 September 1981; accepted 19 January 1982.

1. Thomas, L. & Bowman, M. R. *J. atmos. terr. Phys.* **34**, 1843–1858 (1972).
2. Shimazaki, T. & Laird, A. R. *J. geophys. Res.* **75**, 3221–3235 (1970).
3. Herman, J. R. *J. geophys. Res.* **84**, 3701–3710 (1979).
4. Miller, D. E. & Ryder, P. *Planet. Space Sci.* **21**, 963–970 (1973).
5. Miller, D. E. *Proc. R. Soc. A* **301**, 57–75 (1967).
6. Rodgers, C. D. *Rev. Geophys. Space Phys.* **14**, 609–614 (1976).
7. Broadfoot, A. L. *Astrophys. J.* **173**, 681–689 (1972).
8. Simon, P. C. *COSPAR Technique Manual Ser. No. 7* (Institut d'Aeronomie Spatiale de Belgique, 1978).
9. Inn, E. C. Y. & Tanaka, Y. *Advances in Chemistry* No. 21, 263 (American Chemical Society, Washington DC, 1959).
10. Vigroux, E. *Ann. Phys. Paris* **8**, 705–762 (1953).
11. Penndorf, R. *J. opt. Soc. Am.* **47**, 1976–1820 (1957).
12. Moe, O. E. & Milone, E. E. *Astrophys. J.* **226**, 301–314 (1978).
13. *COSPAR International Reference Atmosphere 1972*, 216 (COSPAR, Akademie, Berlin, 1972).
14. Krueger, A. J. & Minzner, R. A. *J. geophys. Res.* **81**, 4477–4481 (1976).
15. Llewellyn, E. J. & Witt, G. *Planet. Space Sci.* **25**, 165–172 (1977).
16. Hays, P. B. & Roble, R. G. *Planet. Space Sci.* **21**, 273–279 (1973).
17. Krueger, A. J. *et al. Phil. Trans. R. Soc. A* **296**, 191–204 (1980).
18. Gille, J. C., Anderson, G. P. & Bailey, P. L. *J. geophys. Res. Lett.* **7**, 525–528 (1980).
19. Hilsenrath, E. *J. Atmos. Sci.* **28**, 295–297 (1971).
20. Anderson, G. P. *et al. in Proc. Quadrennial int. Ozone Symp.* (ed. London, J.) 580–585 (International Ozone Commission (IAMAP), Boulder, 1981).
21. Penfield, H. *et al. J. geophys. Res.* **81**, 6115–6120 (1976).
22. Wilson, W. J. & Schwartz, P. R. *J. geophys. Res.* **86**, 7385–7388 (1981).
23. Hunt, B. G. *J. Atmos. terr. Phys.* **35**, 1755–1798 (1973).
24. Fabian, P., Pyle, J. A. & Wells, R. J. *J. geophys. Res.* (in the press).
25. *Chemical Kinetic and Photochemical Data for Use in Stratospheric Modelling* Jet Propulsion Laboratory, Pasadena, 1981).
26. Nicolet, M. & Peetermans, W. *Planet. Space Sci.* **28**, 85–103 (1980).
27. Nicolet, M. *Etudes des Reactions Chimiques de l'Ozone dans la Stratosphere* 486–497 (Institut Royal Meteorologique de Belgique, Brussels, 1980); *Planet. Space Sci.* **29**, 951–974 (1981).
28. Frederick, J. E. & Hudson, R. D. *J. Atmos. Sci.* **37**, 1088–1098 (1980).
29. Rogers, J. W. *et al. Geophys. Res. Lett.* **4**, 366–368 (1977).
30. Radford, H. E. *et al. J. geophys. Res.* **82**, 472–478 (1977).

'Paradoxical' mechanics under fluid flow

J. M. T. Thompson

Department of Civil Engineering, University College London, Gower Street, London WC1E 6BT, UK

Interest in the fluid loading of engineering structures has been stimulated by acute technological problems arising in the exploitation of off-shore oil fields. Here slender riser pipes, and complete tethered-buoyant platforms must be designed against potentially dangerous instabilities. We show here how conventional structural theorems can be not only violated, but actually reversed under fluid loading due to its essentially non-conservative character. Under increasing flow, for example, the stiffness of an elastic structure can increase, pass through infinity, and become negative. Surprisingly, the negative stiffness domain is stable, but can be destabilized by the addition of a constraint. Experiments on a hanging articulated pipe conveying fluid nicely confirm the theory.

Suppose we are given a black box with a tray hanging beneath it. We add a small mass m to the tray, and it rises a small distance h ; we add a second mass m and it rises a further distance h . To our surprise, we find that we can plot a complete stable load-deflection curve with negative stiffness using dead weight loading. What is in the box?

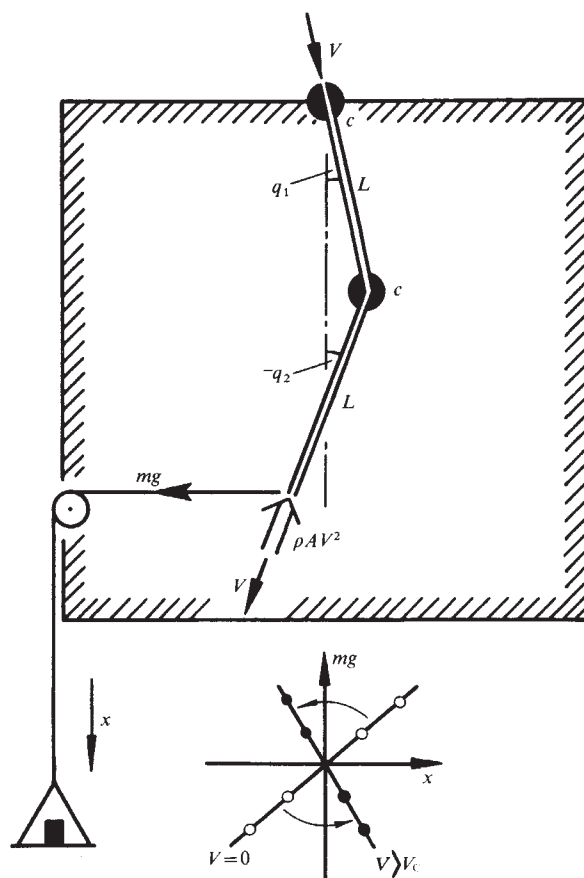


Fig. 1 Articulated pipe system in a box, exhibiting stable negative stiffness under dead loading. The specific details of the system relate to the theoretical study and are not essential. The lengths and stiffnesses need not be the same, and the phenomenon can indeed be exhibited by a continuously flexible hanging hose pipe.

Now a local negative stiffness is quite common in structural mechanics, being, for example, generated by an elastic structure that has lost its initial stability at a fold or limit point: but it is then always associated with unstable equilibrium state under dead load¹. Indeed, an inverted rigid pendulum in our black box would give an equilibrium path of negative stiffness, but the path would be unstable under dead load: it could only be stabilized and therefore studied experimentally by using a suitable rigid-screw loading device.

We can be quite sure, by a fundamental theorem of structural mechanics¹, that our mysterious black box certainly does not contain a conservative mechanical system.

The answer is that it contains an articulated pipe carrying an internal fluid flow, as studied by Benjamin². Two straight and rigid pipes of length L are pinned elastically to one another and to a fixed support by rotational springs of stiffness c , forming a sort of cantilevered hose pipe, the fluid discharging with velocity V into the atmosphere at the free tip.

The statics of the non-conservative system are no mystery. Ignoring gravity, which is inessential in our present theoretical demonstration, the centrifugal force at the knee is statically equivalent to a follower force of magnitude ρAV^2 at the discharging tip as shown in Fig. 1. Here ρ is the density of the fluid and A is the internal cross-sectional area of the pipes. So for small deflections from the vertical, the linearized equilibrium equations are easily written down in terms of the small angles q_1 and q_2 . Moments about the knee for the lower link give

$$mgL + c(q_2 - q_1) = 0$$

and moments about the support for the whole system give

$$2mgL + cq_1 + FL(q_2 - q_1) = 0$$

while the corresponding deflection of mg is

$$x = -L(q_1 + q_2)$$

Here g is the acceleration due to gravity, and $F = \rho AV^2$. Solving these we find

$$x = -mgL^2(2\rho AV^2L/c - 5)/c$$

so the flexibility x/m decreases linearly with V^2 passing from positive to negative as V passes through its critical value V_C

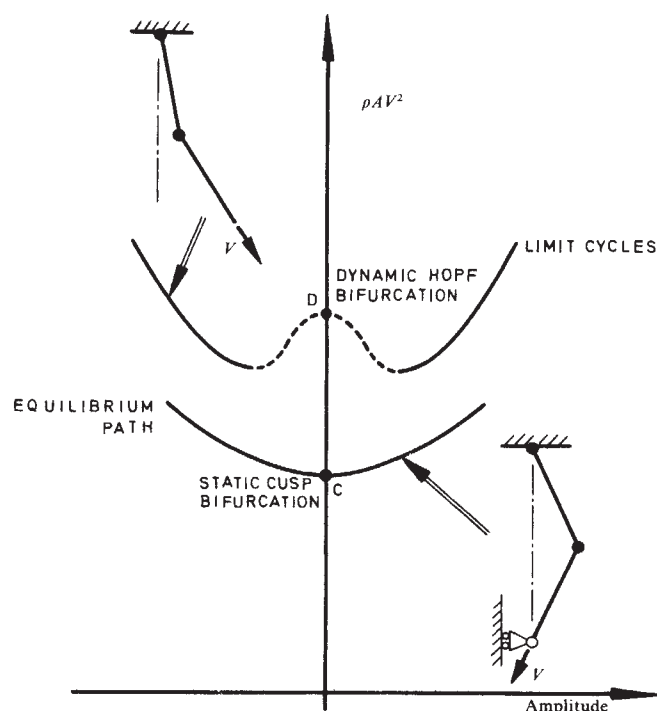


Fig. 2 Buckling (divergence) of a constrained system before the flutter of the corresponding unconstrained system. The dynamic flutter bifurcation is locally sub-critical, but the emerging unstable limit cycles are subsequently stabilized at large deflections.

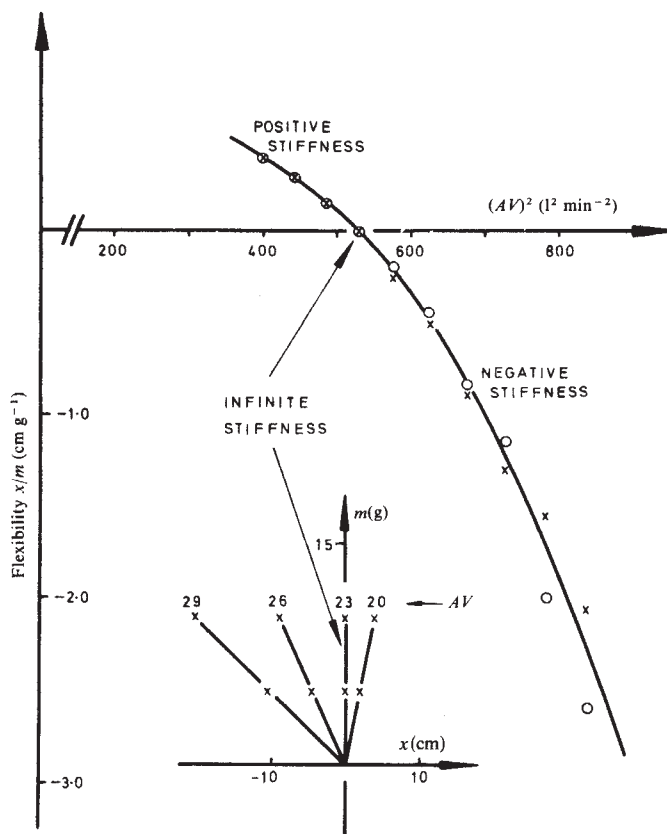


Fig. 3 Experimental results showing a region of stable negative stiffness under dead loading. Infinite stiffness occurs at a critical value of the flow velocity V .

given by setting the expression in brackets to zero. The stiffness m/x therefore increases with V , passes through infinity at $V = V_C$ and becomes negative for higher values of V . In all of this, it is to be understood that V is held constant during the stiffness measurements.

So provided the flow velocity is greater than V_C , we have our negative stiffness. The statics is thus easily understood, but why is the system stable at these high velocities? The answer lies in a complete dynamical stability analysis following Benjamin². The displacement-dependent 'follower' force F at the tip is a circulatory force³, being non-conservative and not derivable from a potential energy function: it is represented by a non-symmetric matrix in the lagrangian equations of motion. In addition, the fluid generates velocity-dependent Coriolis forces.

The dynamical study shows that the system is indeed stable for flow velocities up to and above V_C , stability finally being lost at a dynamic instability at a higher velocity V_D . This dynamic bifurcation to flutter is of the Hopf type, and on the basis of Lunn's new experiments⁴, the emerging limit cycles are seen to be unstable in the vicinity of the branch, with the form of Fig. 2.

The paradoxical static behaviour arises before the onset of this dynamic instability for $V_C < V < V_D$. It is related to the fact, noted by Benjamin², that if the free end were pinned (but free to move vertically) the resulting single degree of freedom system would buckle (diverge) statically at a flow velocity of $V = V_C$. This single degree of freedom pinned system exhibits a stable-symmetric point of bifurcation (stable cusp) at V_C as shown in Fig. 2, analogous to the behaviour of a continuous pin-ended pipe⁵.

Note here a second paradox. When $V_C < V < V_D$ the addition of a constraint ($q_1 + q_2 = 0$) destabilizes the system. This cannot occur in a conservative mechanical system. It was observed experimentally by Benjamin² and repeated by Lunn⁴; if the finger is pressed very lightly against the side of the discharging

tip, the pipes buckle like the pinned system pressing strongly and unexpectedly against the finger.

This brings us to the final paradox. The finger is acting like a semi-rigid screw loading device, under which the system is unstable. In the same flow conditions the pipes are stable under dead loading. This is the exact converse of the behaviour of a buckled elastic structure which is stabilized by semi-rigid loading¹.

Experiments confirming these theoretical deductions are shown in Fig. 3. These used Lunn's experimental pipes, and show the flexibility decreasing with the square of the flow velocity, although not quite in the linear manner of our simple theory which neglected gravitational effects. Perfectly stable equilibrium paths of negative stiffness were obtained in dead-loading conditions. The infinite stiffness at $AV = 23 \text{ l min}^{-1}$ is achieved by the structure buckling progressively in its single degree of freedom mode for which $q_1 + q_2 = 0$, thereby giving the reaction ρAV^2 greater and greater leverage about the top support. It is, of course, eventually destroyed by large deflection nonlinear effects.

This new concept of stable negative stiffness explains the strange behaviour discussed qualitatively by Benjamin, and has important implications for the stability and strength analysis of the flexible marine structures now being proposed.

Received 26 November 1981; accepted 20 January 1982.

1. Thompson, J. M. T. *Phil. Trans. R. Soc. A* **292**, 1–23 (1979).
2. Benjamin, T. B. *Proc. R. Soc. A* **261**, 457–486; 487–499 (1961).
3. Thompson, J. M. T. *Instabilities and Catastrophes in Science and Engineering* (Wiley, Chichester, 1982).
4. Lunn, T. S. thesis, Univ. London (1982).
5. Thompson, J. M. T. & Lunn, T. S. *J. Sound Vib.* **77**, 127–132 (1981).

Structural effects in electrocatalysis

R. R. Adžić & A. V. Tripković

Institute of Electrochemistry, ICTM and Centre for Multidisciplinary Studies, University of Belgrade, PO Box 815, Belgrade, Yugoslavia

W. E. O'Grady

Brookhaven National Laboratory, Department of Energy and Environment, Upton, New York 11973, USA

Studies of chemical reactions on well-defined surfaces provide fundamental data on surface reactivity and guidelines for the understanding and design of catalytic materials but few such studies have been done in electrocatalysis. We report here a study of the oxidation kinetics of HCOOH, CH₃OH and CH₂O on single crystal platinum electrodes with (100), (110) and (111) orientations. Pronounced dependence of the kinetics of these reactions on the crystallographic orientation of the surface has been found. The potential regions where reactions take place, the peaks of voltammetry curves, and the magnitude of currents at the peaks are different for each plane. These differences are explained on the basis of adsorption of a strongly bound intermediate, which shows a pronounced dependence on the symmetry of the single crystal planes. This intermediate completely blocks the Pt (100) surface; the smallest adsorption is at the (111) plane. On activation, the (100) surface shows the highest activity. These results suggest that electrocatalytic reactions exhibit structural sensitivity. They also provide guidelines for designing catalysts for the oxidation of small organic molecules to be used in electrochemical energy conversion.

Kinetics of typical electrocatalytic reactions have been studied on poorly characterized high-area supported electrodes and polycrystalline electrodes. Only a few electrosorption reactions have been studied on single crystal surfaces which include hydrogen adsorption^{1–4} and ad-atom formation (underpotential

deposition of metals)⁵. Previous work on the oxidation of HCOOH, CH₂O and CH₃OH has indicated that these reactions are 'site-demanding' reactions, requiring specific lattice geometries. Bagotzkii *et al.*⁶ showed that in the oxidation of CH₃OH three sites are engaged in adsorption of the intermediate COH. Such reactions will be expected to be sensitive to the structure on (111), (110) and (100) surfaces. The role of specific multiple site requirements was introduced by Balandin⁷ as early as 1929.

The method of preparation and characterization of single crystal surfaces have been described elsewhere^{8,9}. The electrodes were prepared from a Pt single crystal obtained from Materials Research Corporation. The crystals were oriented by Laue back scattering to better than 0.5°. The samples were cut and polished using standard metallurgical techniques and were then slightly etched in boiling aqua regia to remove the work hardened surface. The samples were then annealed at 900 °C for 3–4 h in pure oxygen, reduced at 900 °C for 30 min in hydrogen and cooled over a period of 12 hours in this same atmosphere. They were mounted in Teflon tubing. Immediately before the electrochemical experiment, the electrodes were exposed to 3M HClO₄ to remove possible contaminants such as CaO; they were then rinsed several times with triply distilled H₂O. The electrolyte was 1M HClO₄ prepared from Merck HClO₄ and H₂O prepared by a pyrocatalytic distillation.

Figure 1a shows voltammetry curves for the oxidation of formic acid for three single crystal orientations of Pt. The curves differ substantially in shape and in the magnitude of the currents associated with various peaks. The sweeps in the anodic scan, that is the curves obtained by increasing the potential, show that the highest activity is obtained with the (111) and lowest with the (100) orientation. At the Pt (111) surface the reaction commences at the most negative potential. At the (100) surface the one peak seen sweeping in an anodic direction is usually ascribed to the oxidation of a strongly bound intermediate which blocks the electrode surface for further oxidation of HCOOH. Its oxidation coincides with the onset of oxide formation (100) plane. After activating the surface in this way, the curve obtained by decreasing the applied voltage shows much higher activity. No such effect is seen with the Pt (111) surface: the sweeps almost retrace themselves. This suggests that very little, if any, of the strongly bound intermediate is formed on the (111) plane. The Pt (110) surface shows an intermediate behaviour. Note that the peak in cathodic direction appears at potentials much more positive than the corresponding peak for the (100) plane. Important information is obtained when these curves are compared with that of a polycrystalline electrode. The peak contributions could be correlated with the contributions of the three low Miller index planes. Figure 1b shows a typical voltammetry curve for the oxidation of HCOOH on a polycrystalline platinum electrode. A comparison of the peak potentials on this curve with those obtained for three single crystal orientations (Fig. 1) shows that the first peak in anodic scan (~0.48 V) is characteristic of the (111) plane, and to a lesser degree, of the (110) surface. The second peak in the anodic scan (~0.9 V) is characteristic of the (100) and (110) surfaces. In the cathodic direction, the first peak at 0.7 V is due to the reaction at the (110) surface, while the shoulder at ~0.4 V is characteristic of the (100) plane.

Figure 1c displays voltammograms obtained with methanol. The oxidation commences at 0.5 V at the (111) plane. At the (110) and (100) surfaces the reaction starts at 0.68 and 0.75 V respectively. The largest current is associated with the peak at 0.8 V on the (100) plane. However, from a practical point of view this oxidation takes place at too high a potential because a cell with such electrode would have too low voltage. In that respect, the (111) plane could be considered as the most active catalyst. Note that CH₃OH affects hydrogen adsorption on Pt to a lesser extent than do CH₂O or HCOOH. These data are at variance with the work of Bagotzkii *et al.*¹², who did not find any significant difference between the activity of these single crystal electrodes.

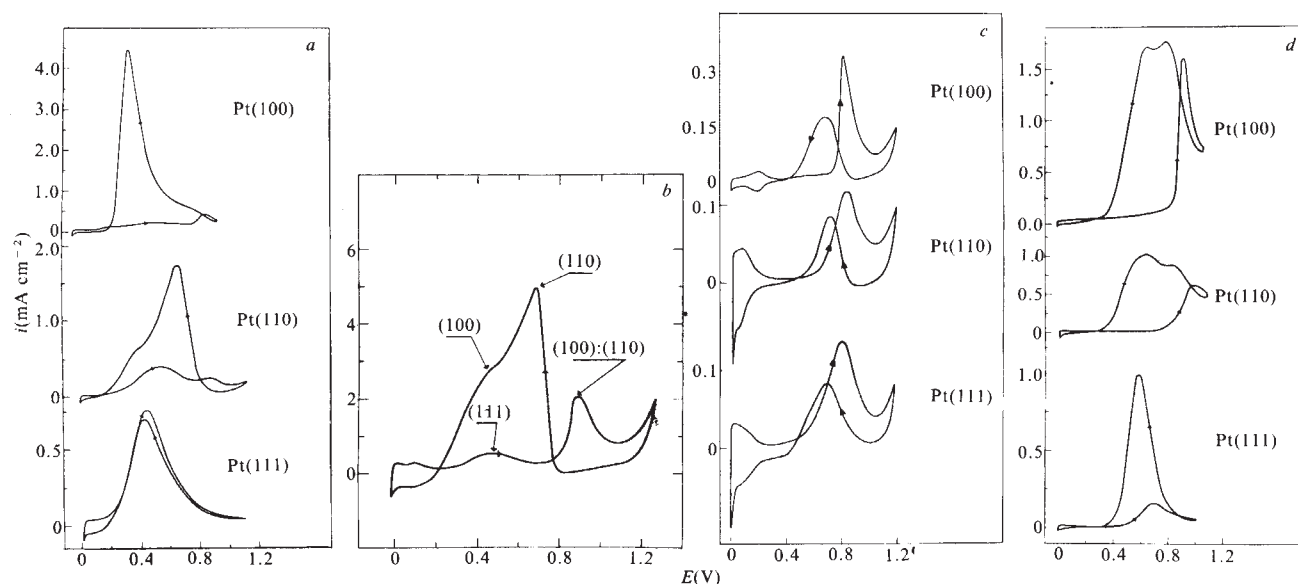


Fig. 1 *a*, Oxidation of HCOOH (0.26 M) on single crystal Pt electrodes with (100), (110) and (111) orientations in 1 M HClO₄. *b*, Oxidation of HCOOH on polycrystalline Pt electrode. Assignments of peaks are given in the graph. *c*, Oxidation of CH₃OH (0.26 M) on single crystal Pt electrodes with (100), (110) and (111) orientations. *d*, Oxidation of CH₂O on single crystal Pt electrodes with (100), (110) and (111) orientations. Sweep rate 50 mV s⁻¹, temperature 22 °C.

The reaction for formaldehyde begins first at the Pt (111) surface, (Fig. 1*d*) at $E = 0.4$ V and reaches a peak at $E = 0.75$ V. At the Pt (100) surface the oxidation commences significantly only at 0.75 V, with the peak at 0.88 V. The current associated with this peak is, however, higher than the current for the (111) and (110) surfaces. After activation of surfaces at the most positive potentials, the sweeps in the cathodic direction show the highest activity on the (100) plane. This shows that strongly bound intermediates in the oxidation of CH₂O on Pt are predominantly formed on the (100) and (110) planes.

It is usually assumed that in the oxidation of HCOOH, CH₃OH and CH₂O, which should lead to CO₂, the same strongly bound COH intermediate is formed which blocks the surface, thus decreasing its catalytic activity. Differing activities of the (111), (110) and (100) Pt single crystal electrodes are ascribed to different degrees of adsorption of COH species. Adsorption is lowest on the (111) surface and highest on (100) plane, with (111) orientation generally exhibiting low reactivity. Low hydrogen adsorption on the Pt (111) surface, found by most authors¹⁻³, also has a role in determining the activity of this surface for the oxidation of HCOOH, because adsorbed H participates in the formation of COH. The adsorption and reduction of CO₂ on Pt, which gives the same species, COH, also gives the smallest coverage on the (111) surface¹³. This corroborates the above conclusion. Crystallographic anisotropies of the above reactions in many ways mimic those found in the gas phase for CO adsorption¹⁴ on the same crystal planes. CO is the species closest in character to COH, which provides the basis for the comparison of these data.

There are no corresponding studies of adsorption and dehydrogenation of HCOOH, CH₃OH and CH₂O on Pt single crystals. As in the case of CO, the anisotropies which we have found cannot be predicted on the basis of theories of adsorption¹⁵ and electrode reactions¹⁶ presently available.

On the basis of these data we conclude that the electrochemical oxidation of HCOOH, CH₃OH and CH₂O on Pt exhibits significant crystallographic anisotropies. As these are typical electrocatalytic reactions, which involve a strong adsorption before the charge transfer, these data strongly suggest that most such reactions should show a pronounced structural sensitivity.

Note added in proof: Since submission of this paper, Clavilier *et al.*¹⁷ have reported similar results.

Received 7 September 1981; accepted 6 January 1982.

- Yeager, E., O'Grady, W. E., Woo, M. Y. C. & Hagans, P. L. *J. electrochem. Soc.* **125**, 348-349 (1978).
- Will, F. G. *J. electrochem. Soc.* **112**, 451-454 (1965).
- Yamamoto, K., Kolb, D. M., Kotz, R. & Lehmpfuhl, G. *J. electroanal. Chem.* **96**, 233-237 (1979).
- Clavilier, J., Faure, R., Guinet, G. & Durand, R. *J. electroanal. Chem.* **107**, 205-209 (1980).
- Bewick, A. & Thomas, B. J. *electroanal. Chem.* **65**, 911-932 (1975).
- Bagotskii, V. S. & Vassiliev, Yu. B. *Electrochim. Acta* **11**, 1439-1448 (1966).
- Balandin, A. A. *Z. phys. Chem.* **B2**, 289-295 (1929).
- Adžić, R. R., O'Grady, W. & Srinivasan, S. *Surface Sci.* **94**, L191-L194 (1980).
- Adžić, R. R., Tripković, A. V. & O'Grady, W. E. *Proc. 3rd Symp. on Electrocatalysis*, Minneapolis (in the press).
- Capon, A. & Parsons, R. *J. electroanal. Chem.* **44**, 239-249 (1973).
- Adžić, R. R. *Israel J. Chem.* **18**, 106-181 (1979).
- Pyshnograeva, I. I., Vasilyev, Yu. B. & Bagotskii, V. S. *Elektrokhimiya* **6**, 433-438 (1970).
- Marković, N. & Adžić, R. R. *Proc. 7th Yugoslav Symp. on Electrochemistry*, Ohrid, 157-161 (1981).
- McCabe, R. W. & Schmidt, L. D. *Surface Sci.* **66**, 101-117 (1977).
- Roberts, M. W. & McKee, C. S. *Chemistry of the Metal-gas Interface* (Oxford University Press, 1978).
- Conway, B. E. *Theory and Principles of Electrode Processes* (Ronald, New York, 1965).
- Clavilier, J. *et al. J. electroanal. Chem.* **124**, 321-324; **125**, 249-254 (1981).

Controlled nucleation and quasi-ordered growth of ice crystals from low temperature electrolyte solutions

J. Dupuy*, J. F. Jal*, C. Ferradou*, P. Chieux†, A. F. Wright†, R. Calemczuk‡ & C. A. Angell§

*Laboratoire de Physique des Matériaux, Université Claude Bernard, Lyon, France

†Institut Laue-Langevin, Avenue des Martyrs, 38042 Grenoble Cedex, France

‡Service des Basses Températures, Laboratoire de Cryophysique, Centre d'Etudes Nucleaires de Grenoble, 85 X, 38041 Grenoble Cedex, France,

§Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA

The process by which ice forms during the cooling of water and aqueous solutions has been studied intensively mostly in conditions where the nucleation occurs heterogeneously—on some particulate impurity surface. When the more fundamental homogeneous process has been deliberately investigated¹⁻³ it has nearly always been in conditions in which the crystallization

process, once initiated, occurs with great rapidity^{1,2}. Seeking an additional 'degree of freedom' in the investigation of this important phenomenon, we have initiated the crystallization in conditions where the growth is extremely slow and the nucleation events can be controlled. Our observations are described below.

Slow nucleation conditions are obtained by choosing a concentration of water in an electrolyte solution such that the homogeneous nucleation temperature, which may be obtained in separate experiments using emulsions^{2,3}, lies very close to the glass transition temperature. In these conditions the growth of fluctuations to become, first, nuclei and then observable crystals, occurs at a rate which may be varied at will by changing the temperature of observation within a narrow range of a few degrees on either side of the glass transition temperature⁴.

The study of nucleation of ice in such conditions has gained in interest since the observations by A.F.W. and colleagues⁵ of the manner in which crystallization patterns of inorganic high temperature glasses can be influenced by the preceding thermal history. It was demonstrated that, for the appropriate heat treatment, crystallites could be observed growing with correlated spatial distributions with crystallization centres as small as 20 Å in radius and mean separations in the range 200–700 Å depending on thermal history. Since the origin of this phenomenon was not clear at the time (1979), it was clearly of interest to determine whether similar ordered crystal growth could be observed in a different type of system. We therefore undertook a careful study of the development and distribution of crystallites of ice formed in an initially vitreous lithium chloride + D₂O solution, using the small angle neutron scattering (SANS) technique. The conditions for glass formation, and observation of diffusion-controlled nucleation, have been described elsewhere⁴.

Solutions of 10.5 mol% lithium chloride in D₂O (6.17 m LiCl) were prepared in controlled conditions to avoid the ingress of any light water, H₂O. Samples of the solutions were introduced into a 3-mm space between sapphire windows in a cell whose temperature could be manipulated by a controlled temperature block in which the cell was mounted. The block with cell was then suspended in a liquid helium cooled cryostat which provided the low temperature ambient for the temperature control system. The temperature of the solution was taken to be that of the sapphire window at the edge of the cell, as preliminary experiments had shown that a gradient of <0.1 °C existed between the centre and edge of the sample. The temperature of the cell was monitored during the experiment by a fine chromel/constantin thermocouple in good thermal contact with the window.

The sample was initially vitrified by quenching into liquid nitrogen placed in the bottom of the cryostat and monitoring the temperature to ensure that it did not fall below 130 K. Under these circumstances no cracks developed in the sample: the sample remained completely transparent to visible light and caused very little scattering of neutrons. (In some cases where the temperature drop was inadequately controlled, and cracks due to mechanical stress developed, the effects (observed as an excess scattering), could be seen to disappear during the first 2 h of annealing at 139 K.)

On the basis of previous observations⁵ of the effect of annealing below T_g on the kinetics and spatial characteristics of nucleation, and with the assistance of preliminary calorimetric and light-scattering studies, it proved possible to determine thermal histories which would produce ordered distributions of microscopic ice crystals in our system. The conditions are critical, variations in choice of annealing temperature >1 °C not being allowed. Spectra showing the peak in Q described below were finally obtained using a nucleation temperature of 139 K (compare with $T_g = 142$ K) for a period of 48 h, followed by observation at 140–142 K during which the nuclei which formed at the lower temperature grew to observable dimensions.

Following the annealing period, the sample still at 139 K was placed in the cold neutron beam of the D11 (low angle scatter-

ing) instrument of the ILL HFR (Grenoble). The wavelength of the study ($\lambda = 9$ Å) and the position of the multidetector (10 m behind the sample) were chosen so as to detect the correlations in the spatial distributions of the nuclei centres. With these conditions we obtained a range of scattered neutron wave vectors $2 \times 10^{-3} < Q < 3 \times 10^{-2}$ Å⁻¹. The counting time, 20 min, was a compromise based on the need to acquire reasonable statistics, while still following the time development of the system at a rate compatible with total allotted beam time.

At 139 K the sample gave only a very weak, featureless and time-independent scattered signal. This remained true while the temperature was raised slowly to 140.4 K. At this temperature, a statistically significant increase in scattering was observed, and the temperature increase was halted to permit isothermal observation of the sample time evolution (see Fig. 1). The scattered intensity which developed after 4 h is shown in Fig. 2a, and is notable for the intensity maximum at $Q_m = 0.9 \times 10^{-2}$ Å⁻¹. This feature indicates some spatial organization of scattering centres, with a mean distance of separation of 700 Å.

After 6 h, the evolution effectively ceased, and to accelerate the development the temperature was raised (see Fig. 1). An increase in intensity and loss of peak definition was observed as the temperature rose to 143 K but thereafter an unexpected decrease in intensity occurred and continued until the signal virtually disappeared at 149 K (see Fig. 2b). The signal then reappeared, with a suggestion of a maximum in Q at the same value as before, and grew rapidly. This second increase was followed isothermally at 151 K (see Fig. 1) for 2 h. According to separate observations the sample, up to this point, remains optically transparent. Finally, observations were made during continuous temperature increase to $T > 153$ K, during which the sample becomes optically opaque, implying the presence of many large particles produced by ripening in the more diffusive conditions. Visible light scattering in the presence of a maximum in Q was also observable in microcrystallized cordierite glasses in the case where $\kappa_{\max} < 1.5 \times 10^{-2}$ Å⁻¹ (ref. 7).

The observation of a maximum in the scattered intensity spectrum following annealing treatments similar to those used previously⁵ for silicate systems suggests that the ordering of crystallite centres has a common origin. The origin has been discussed elsewhere⁶ in terms of stabilization of crystallization centres initiated in highly metastable conditions in which the nucleation probability is high (because the critical nucleus radius is very small) and the diffusion length is small. The latter condition allows particles to grow independently of one another despite small separations if they lie beyond the region of steep concentration gradient established around each growing particle. New nucleation in the region of influence of each established

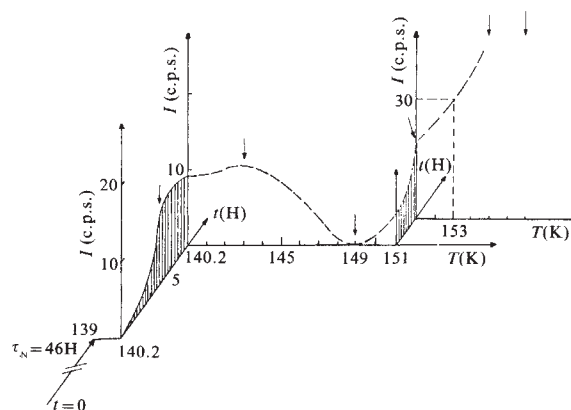


Fig. 1 Time-temperature-scattered neutron intensity schematic for this experiment. Note particularly the long annealing time at 139 K, the gradual development of scattered intensity with time at 140.2 K and the relatively sudden loss of intensity in the range 143–149 K. Short down-pointing arrows indicate conditions for the spectra shown in Fig. 2.

particle quickly becomes improbable in comparison with further growth of the established particle. Thus a preferred distance of separation, which is smaller the lower the temperature of nucleation, becomes established, and resists dissipation during the subsequent temperature increase to the value chosen for observation.

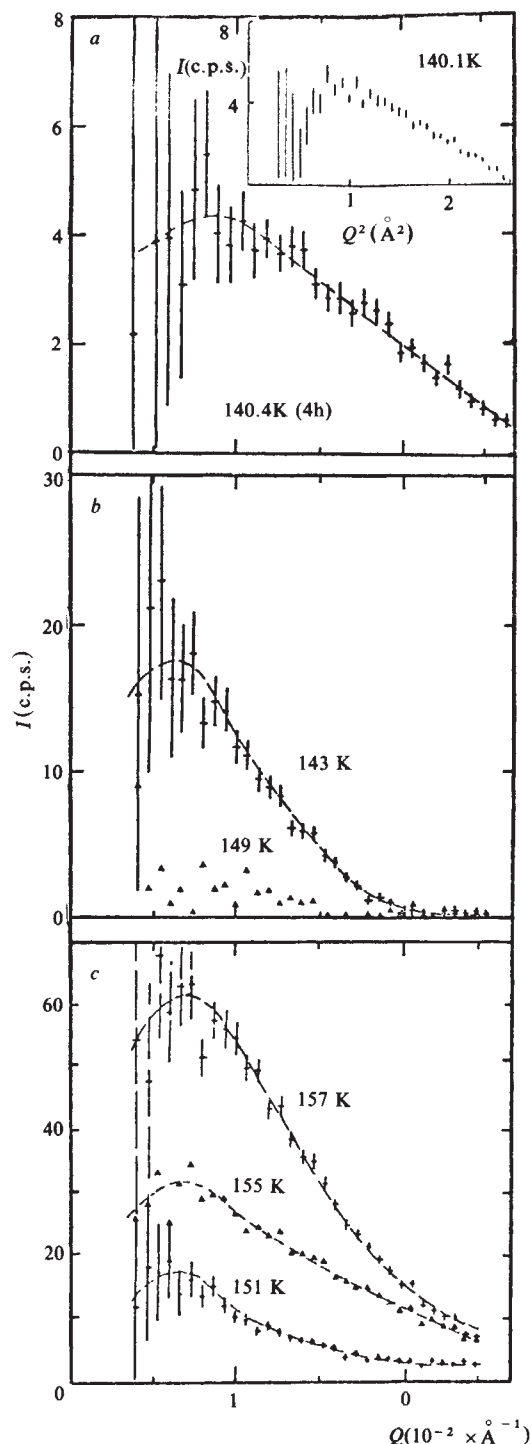


Fig. 2 Scattered neutron intensity in c.p.s. as a function of Q^2 at different stages of the experiment. *a*, After 4 h of development of 140.2 K. *b*, During temperature change in the range 143–149 K showing disappearance of signal. *c*, Development of signal with temperature increase between 151 K and 157 K. Note the maximum in I at $Q^2 \approx 0.9 \times 10^{-2} \text{ \AA}^{-2}$ in 140.4 K spectrum which persists throughout the experiment, with some movement to lower Q^2 values at higher temperatures. The same feature was seen better resolved (see insert to *a*) in an earlier exploratory experiment, for which the time–temperature history in the annealing range was less simple.

Considering the mean separation of crystallites of 700 Å, we observe that the nucleation density at 139 K must be at least 2,500 centres per μm^3 —so high that the homogeneous nature of the nucleation phenomenon under study cannot be doubted. A Guinier plot of the high Q side of the peak yields an average crystallite size for the small particles, after initial development at 140.4 K, of ~ 100 Å. We can confirm approximately the size of particles which must be present by dividing the total ice volume which would form at complete desaturation (determined from the metastable extension of the ice saturation line in the phase diagram)⁴ by the number of scattering centres in the sample volume containing the precipitated ice. This calculation yields a maximum particle radius of 220 Å. Because our observation is made early during the desaturation process, an average radius of 100 Å seems reasonable.

The loss of scattering intensity at intermediate temperatures at first suggested that the initial scattering might have been due to a liquid–liquid phase separation which, as in the case of lithium disilicate⁷, produces conditions favourable for crystal nucleation. The new scattering centres which form in these conditions at a temperature above 149 K should produce a SANS spectrum unrelated to the initial pattern produced by the phase separation. That the renewed scattering pattern in the present case apparently retained the same maximum in Q as the initial scattering pattern suggests a more mundane explanation.

Neutron diffraction⁷ and calorimetric studies⁸ have now shown that the SANS phenomena observed result from a continuous development of microcrystals for $T > 140$ K. This suggests that the peculiar variation in the scattered intensity for 143 K $< T < 149$ K (Fig. 1) probably originates in the relative variations in density and molar scattering power of the precipitated ice and the residual solution, which we can only assess approximately.

Combining data for the various isotopes, we find the scattering cross-section per mole of solution decreases rapidly with increasing concentration of LiCl. If, due to its large molar volume, the scattering power of ice initially formed from the supersaturated solution is weaker than that of the initial solution then, as the ice precipitates and residual solution concentrates, the contrast must pass through zero, giving a scattered intensity minimum. The low temperature density data for LiCl+D₂O solutions and D₂O ice necessary to confirm this hypothesis are not available, but extrapolations of LiCl+H₂O solution data, available down to -60°C (ref. 9), show the postulated crossover is entirely possible. The explanation is also consistent with the relatively weak scattering observed at all temperatures.

Finally, the type of measurement described here provides a valuable basis for studies of other physical properties which, while sometimes more sensitive to the formation of the crystallizing phase⁸, contain no direct information on its size or distribution. These experiments and initial studies carried out at the eutectic composition, have produced no evidence in favour of a liquid–liquid phase separation occurring near T_g , either at the eutectic composition¹⁰ or below¹¹.

C.A.A. thanks the NSF for support of this work during a sabbatical leave from Purdue University at the Institut Laue-Langevin, France, under grant no. DMR77-04318A1. We thank D. Brochier and S. Pujol for help in the design and fabrication of the sample temperature control block, and J. P. Lelieur for preliminary calorimetric measurements.

Received 16 November; accepted 29 December 1981.

1. Wood, G. R. & Walton, A. G. *J. appl. Phys.* **41**, 3027 (1970).
2. Rasmussen, D. H. & MacKenzie, A. P. *Water Structure at the Water Polymer Interface*, ed. Jellinek, H. H. G. 126–145 (Plenum, New York, 1972).
3. Kanno, H. & Angell, C. A. *J. phys. Chem.* **81**, 2639 (1977).
4. Angell, C. A., Sare, E. J., Donnell, J. & MacFarlane, D. R. *J. phys. Chem.* **85**, 1461 (1981).
5. Wright, A. F., Talbot, J. & Fender, B. E. F. *Nature* **227**, 366 (1979).
6. Wright, A. F. in *Symp. on Neutron Scattering*, Argonne (American Institute of Physics Symp. Ser., in the press).
7. Elarby, A. *et al. J. Phys. Lett.* (submitted).
8. Angell, C. A. & MacFarlane, D. R. in *Advances in Ceramics* (in the press).
9. Bressel, R. D. thesis, Purdue Univ. (1972).
10. Hsieh, S. Y., Gammon, R. W. & Montrose, C. J. *J. chem. Phys.* **56**, 1666 (1972).
11. Angell, C. A. & Sare, E. J. *J. chem. Phys.* **49**, 4714 (1968).

Setting of gabbroic dykelets in an ophiolitic complex by hydraulic fracturing

I. Reuber, H. Whitechurch & J. M. Caron

Institut de Géologie, 1 rue Blessig, 67084 Strasbourg Cedex, France

Gabbroic dykelets cutting peridotitic tectonites are generally interpreted as the product of *in situ* partial melting, or as the path of upward moving magma. However, we report here that their structure strongly suggests emplacement by hydraulic fracturing, and their bulk and mineral chemistry indicate that they were highly differentiated liquids issued from the overlying magma chamber. They cannot be the feeding channels of the magma chamber, which contains the primitive liquids produced by partial melting of the upper mantle.

Within ophiolitic complexes, several types of dykelet (usually 1–5 cm wide, exceptionally up to 50 cm) can be observed in both the tectonite and the cumulate sequences. They are called dykelets to distinguish them from the isolated diabase dykes¹, which are ~2 m wide and cut throughout the ophiolitic sequence. Because the small dykelets have previously been interpreted as the feeding channels of the cumulate magma chamber^{1,2}, in looking for the path and mechanism of feeding for the chamber we concentrated our investigations on the Antalya ophiolite complex (Fig. 1), which is particularly rich in internal structures related to ridge activity^{3,4}.

This ophiolite consists of two well-exposed complexes (Fig. 1), the tectonic harzburgite massif of Adrasan in the south, and the massif of Cirali/Tekirova ~15 km farther north, consisting of harzburgitic tectonites and magmatic ultrabasic and basic cumulates. The coarse-grained texture ('S1') of the tectonites is due to high temperature, high-pressure, low-creep stress plastic flow below an oceanic ridge. However, in Adrasan and at the south end of the Cirali/Tekirova massif, shear zones with a porphyroclastic texture ('S2') are evidence of a low temperature, low lithostatic pressure, high creep stress deformation, obviously linked to intra-oceanic thrusting before obduction^{5–8}.

A group of leucocratic gabbro dykelets can be observed among all dykelets occurring either in the tectonites (commonly deformed dunite and pyroxenite veinlets) or in the cumulates (nondeformed pyroxenite and gabbro dykelets) (I.R., in preparation). They never exceed 1% of the total rock, even in areas where they appear to be abundant, and may total ~0.05% of the massif. Their abundance in the tectonites decreases rapidly downwards in the upper 500 m.

These gabbroic dykelets neatly cross-cut all structures due to and affected by the plastic flow below the ridge, as well as the magmatic structures in the cumulates. Their mineralogy and chemistry are similar in both parts of the ophiolite, and quite different from those of the other dykelets. In fact, their bulk rock as well as their mineral chemistry (bulk rock chemistry analyses are representative of the total dykelet, as the sharp contact between dykelet and host rock permitted easy separation of the two, Figs 2, 3) correspond to those of evolved residual magmatic liquids, similar to the liquids which crystallized in residual pockets (Fig. 4e) in the upper part of the cumulate sequence. As Fig. 2 shows, there is a marked decrease in the MgO content and an increase in the Al₂O₃ content for the same SiO₂ and CaO contents, compared with those of the host plagioclase-wherlite or host gabbros of the cumulate sequence. The most differentiated dykelets are those found within the harzburgitic tectonites. This same trend occurs in the chemistry of clinopyroxenes and orthopyroxenes (Fig. 3), which show an enrichment in Fe content. Again, this enrichment is most important for the dykelets in the tectonites.

Thus, the leucocratic gabbro dykelets could not be the feeding channels of the overlying cumulate magma chamber. Their chemistry does not correspond to that of liquids produced by upper mantle partial fusion, nor to a possibly picritic parental magma, from which a thick layer of ultramafic cumulates could have crystallized, but it does correspond to the composition of residual liquids developed in the overlying gabbroic cumulates. Actually, the composition of MORB primitive liquids^{9–12} corresponds to a mean value for the cumulates of Antalya (except for CaO values, which are quite dispersed anyway). However, the composition of the gabbro dykelets is considerably more evolved. Neither do they resemble that of the Cyprus upper pillow lavas¹³, believed to represent off-axis magmas¹⁴. When the Omani gabbro dykelets, partly considered to be off-axis magma⁸, are replotted, we find that they can be divided into two groups: Al₂O₃-poor, MgO-rich dykelets, which correspond to little fractionated cumulate composition, and Al₂O₃-rich, MgO-poor dykelets, which might be the nondeformed ones; their

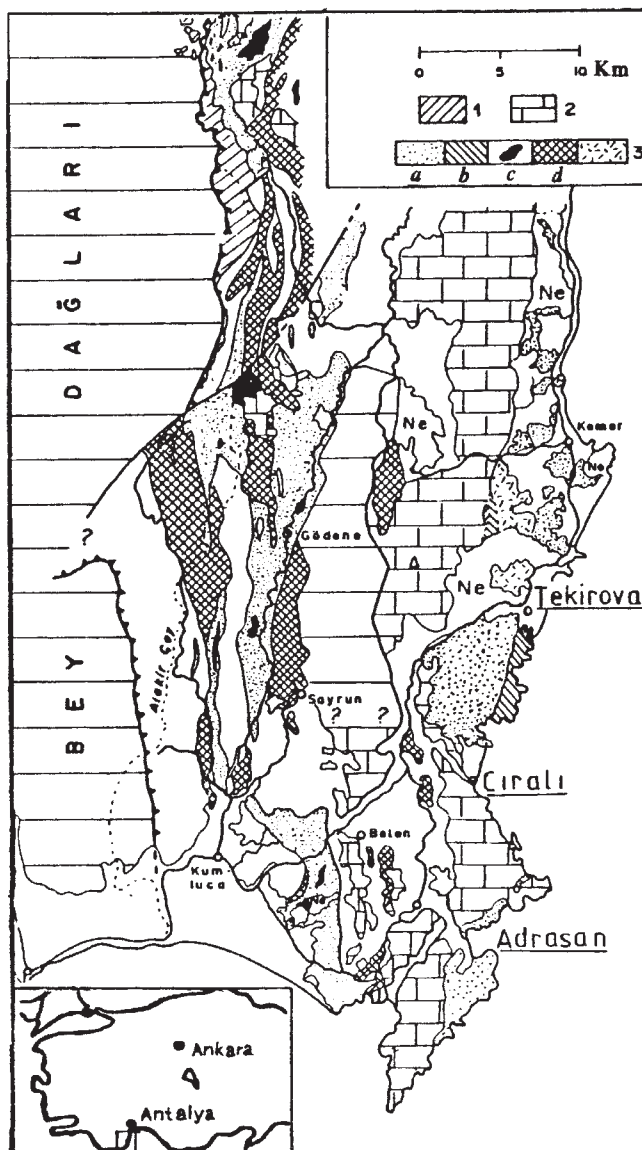


Fig. 1 Simplified structural and petrographical map of the study area, the southwestern segment of the Antalya Nappes, showing the distribution of ophiolitic facies in the median nappe and the two massifs of main interest: the tectonite massif of Adrasan and the tectonite/cumulate massif of Cirali/Tekirova. (1) Lower Antalya Nappe; (2) Upper Antalya Nappe; (3) Median Antalya Nappe, with: a, tectonites (mainly harzburgites and dunites, more or less serpentinized); b, layered ultramafic-mafic cumulates; c, non-layered gabbroic cumulates; d, pillow lavas (upper Trias); e, dyke complex of Kemer. White areas: sedimentary unit of Alakir Cay. Ne: Neogene cover.

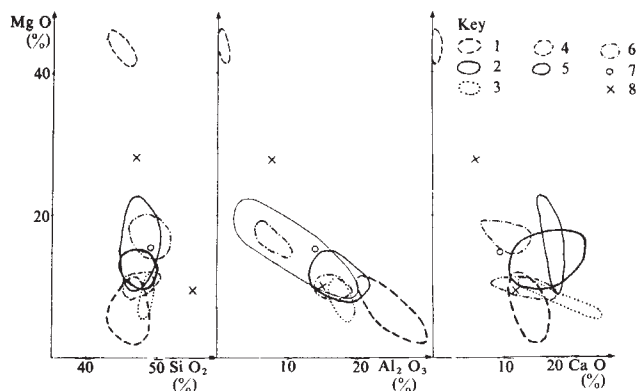


Fig. 2 Harker diagrams of bulk rock chemistry (27 analyses) showing that the differentiation from cumulate host rock—residual pockets and gabbro dykelets within cumulates—gabbro dykelets within tectonites mainly consists of a decrease in MgO and an increase in Al_2O_3 . 1, Gabbro dykelets in the tectonites; 2, gabbro dykelets in the cumulates; 3, residual pockets in the cumulates; 4, tectonite host rock; 5, cumulate host rock; 6, gabbro dykelets from Oman⁸; 7, MORB primitive liquid (mean of values from refs 9–12); 8, Cyprus upper pillow lavas¹³.

emplacement might be explained in the same way as for Antalya.

Thus, if the leucocratic gabbro dykelets within the harzburgitic tectonites are residual liquids coming from the overlying magma chamber, what is the mechanism of downward injection? Their geometry, although varying from one outcrop to another, strongly suggests their setting by hydraulic fracturing (see ref. 2): (1) The end of the dykelets may be marked by small, branching and zigzagging dykelets, matching the images found by Beach¹⁵ for experimental hydraulic fractures. (2) At the scale of a single gabbro dykelet, anastomosing dykelets and magmatic brecciation are frequent at the dyke walls (Fig. 4a). (3) Such brecciation may result in interconnected anastomosing dykelets (Fig. 4b), and even in large brecciated areas, occurring preferentially near the contact between tectonites and cumulates. As the liquid is more abundant, they are better developed on the cumulate side (Fig. 4c). (4) Straight, long, parallel dykelets have also been set by hydraulic fracturing (see below) (Fig. 4d).

Hydraulic fracturing probably occurred after the interstitial liquid had been expelled from between the cumulus minerals, since traces of such interstitial minerals are rarely observed in

the Antalya cumulates. This interstitial liquid would be gathered in residual pockets, from which extending dykelets are frequently observed (Fig. 4e), indicating that liquid overpressure within the pockets provoked fracturing. Relative abundance of liquid (locally exceeding 20% of the total rock) facilitates brecciation, particularly affecting the less permeable rocks, such as the dunites at the contact between tectonites and cumulates.

The variable geometry of these gabbroic dykelets could be explained by the interference between stresses of different origin², that is, fluid pressure (p), deviatoric stress ($\Delta\sigma$) and confining lithostatic pressure. Cracks created by hydraulic fracturing, and thus directly dependent on the fluid pressure, will be found in conditions intermediate between the following cases: (1) Deviatoric stress is zero: fluid pressure need only overcome the confining pressure and the rock coherence. As both pressures are isotropic, cracks should be created equally in all directions, unless guided by rock anisotropy, and the result will be anastomosing veins. (2) Deviatoric stress is medium: cracks will be guided by the anisotropic stress field. As the confining pressure is weak, tensile failure parallel to $\sigma_1 \times \sigma_2$ plane is

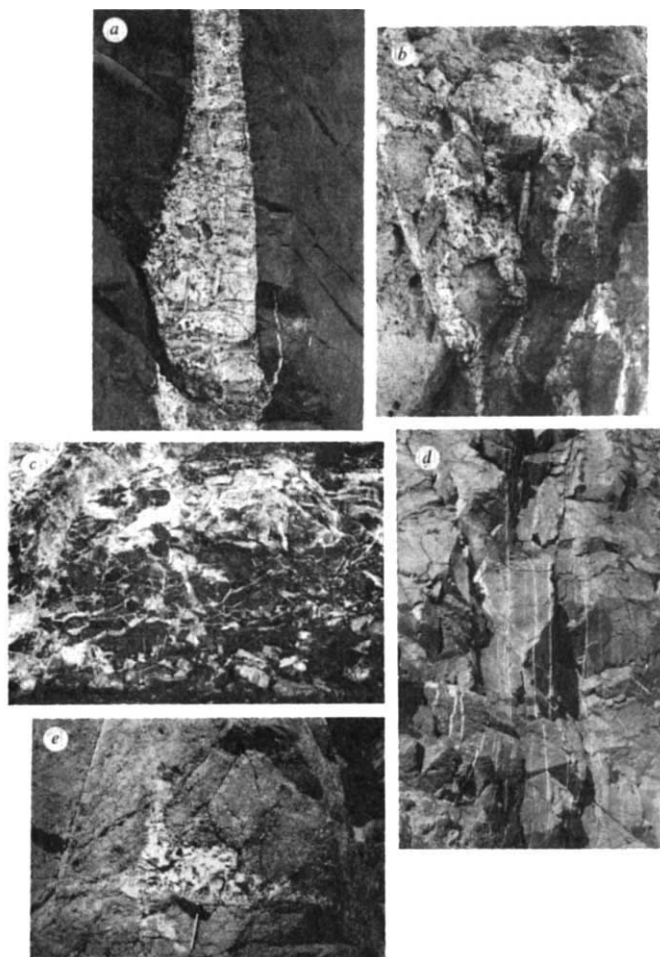


Fig. 4 a, Leucocratic gabbro dykelet (10–25 cm wide) within the tectonites of Adrasan. Foliation S2 is parallel to the dykelet and especially strong in the mylonitic shear zones. The small, branching dykelet is typical of hydraulic fracturing. b, Anastomosing gabbro dykelets, resulting in brecciation of the harzburgitic host rock, tectonite massif of Adrasan. Marker (centre) is 14 cm long. c, Brecciation at a larger scale: the 'megabreccia' of ultramafic cumulates (elements are ~1 m in diameter), cemented by gabbro. Massif of Cirali/Tekirova. d, Straight, parallel dykelets (1–5 cm) in the harzburgites of the Adrasan tectonite massif. e, A pegmatitic residual pocket with an amphibolitic core in melanocratic gabbro of the Cirali/Tekirova cumulates. Note the leucocratic dykelets extending in all four directions. Pencil (bottom centre) is 15 cm long.

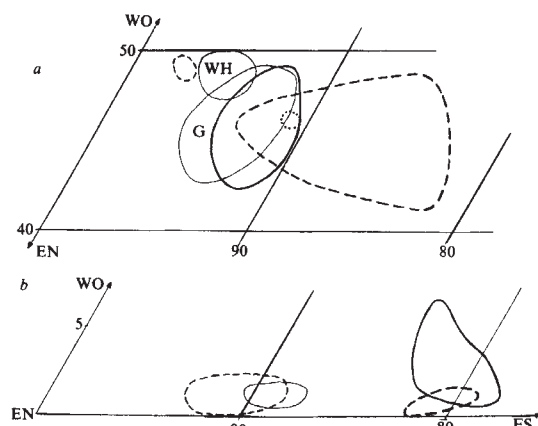


Fig. 3 Plot of clinopyroxene (a) (40 analyses) and orthopyroxene (b) (22 analyses) composition in a greatly enlarged En-Fs-Wo triangle, showing two groups neatly separated for the orthopyroxene: host rock orthopyroxene at En_{90} , and orthopyroxene of the dykelets clearly richer in Fe with En_{80} . With corresponding values the clinopyroxene show a continuous enrichment in Fe from tectonite host rock, via cumulate host rock (WH = wherlite, G = gabbro) and gabbro dykelets in the cumulates, to gabbro dykelets in the tectonites. Key as for Fig. 2.

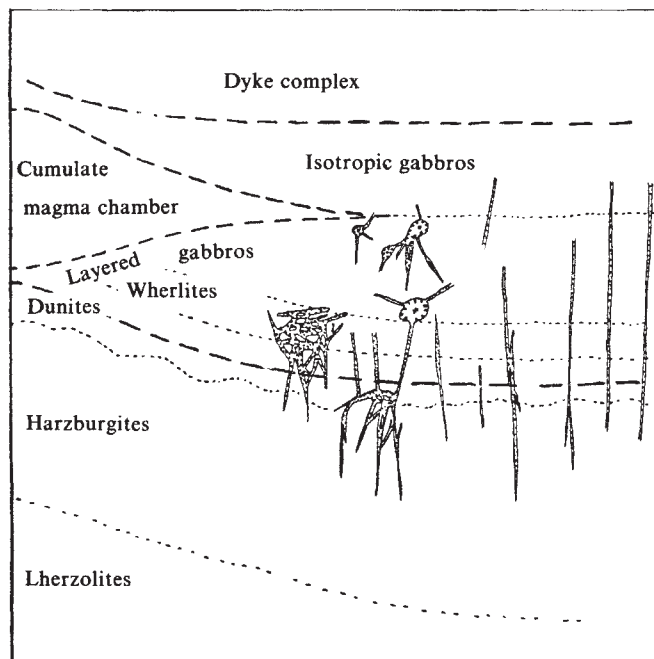


Fig. 5 Schematic model of a mid-oceanic ridge to show that the emplacement of the leucocratic gabbro dykelets by hydraulic fracturing takes place at a late stage in the evolution of the cumulate magma chamber.

possible, if the fluid pressure p yields $p > 3\sigma_2 - \sigma_1 + T_0$ (where T_0 = tensile strength¹⁶). Dykelets resulting from such tensile fracturing initiated by fluid pressure will be long, straight, and parallel to one another. (3) Deviatoric stress is high: cracks will still be guided by the anisotropic stress field, but instead of tensile fractures, shear fractures will develop. Such fractures are oblique on the $\sigma_1 \times \sigma_2$ plane, short and irregular, with a characteristic *en echelon* arrangement. With increasing fluid pressure the propagation speed of such cracks, and its deviation from the theoretical shear fracture direction, will increase¹⁷. The resulting more irregular fractures and especially their *en echelon* arrangement have also been observed in a lesser amount. They appear to be anterior to the anastomosing and then to the straight dykelets.

The stress conditions in an oceanic spreading centre clearly depend on the locality below the ridge. Thus, a given rock assemblage will be submitted to varying stress regimes as it travels away from the ridge centre. Hydraulic fracturing, although occurring at a late stage with respect to the cumulate magma chamber's evolution, must be active over a longer span of time, as several generations of dykelets produced by hydraulic fracturing, or dykelets cutting magmatic breccias are observed. At one outcrop the younger the dykelet, the more differentiated its chemistry. In general the succession of the different vein geometries points to high deviatoric stress conditions nearest to the spreading centre, passing via isotropic conditions to tensile stress away from the ridge.

In conclusion, the leucocratic gabbro dykelets cutting the tectonite and cumulate sequences of the Antalya ophiolitic complex, constitute the latest generations of all the dykelets observed (except for the much later injected diabase dykes). They have been formed by hydraulic fracturing in variable stress conditions. They are not the feeding channels of the overlying magma chamber, which must be looked for among the other dykelets (I.R., in preparation), but their very differentiated chemistry indicates that they have originated from the overlying cumulates. They have been injected at a late stage of the cumulate magma chamber evolution in all directions in the cumulates, and downwards into the tectonites (see Fig. 5).

We thank A. Nicolas, R. G. Coleman and T. Juteau for helpful remarks and discussions. Financial support was provided by ATP 'Geodynamique' from the French INAG.

Received 26 May; accepted 18 December 1981.

1. Juteau, T., Whitechurch, H. *Proc. int. Ophiolite Symp.*, Cyprus, 378–391 (ed. Panatoyou, A.) (1980).
2. Jackson, M. thesis, Nantes Univ. (1979).
3. Juteau, T. *Sci. Terre, Mem.* **32** (1975).
4. Juteau, T., Nicolas, A., Fruchard, J. C. & Bouchez, J. L. *Bull. geol. Soc. Am.* **88**, 1740–1748 (1977).
5. Parrot, J. F. & Whitechurch, H. *Rev. Geogr. Phys. Geol. Dyn.* **20**, 153–170 (1978).
6. Cakir, Ü., Juteau, T. & Whitechurch, H. *Bull. Soc. géol. Fr.* **20**, 61–70 (1978).
7. Nicolas, A. & Le Pichon, X. *Earth planet. Sci. Lett.* **46**, 397–406 (1980).
8. Boudier, F. & Coleman, R. G. *J. geophys. Res.* **86**, 2573–2592 (1981).
9. Irvine, T. N. *Yb. Carnegie Instn Wash.* **76**, 454–461 (1977).
10. Irvine, T. N. in *The Evolution of Igneous Rocks* (ed. Yoder, H. S.) 245–306 (Princeton University Press, 1979).
11. Elton, D. *Nature* **278**, 514–518 (1979).
12. Laurent, R., Delaloye, M., Vagnat, M. & Wagner, J. S. in *Proc. int. Ophiolite Symp.*, Cyprus, 172–181 (ed. Panatoyou, A.) (1980).
13. Desmet, A. thesis, Nancy Univ. (1977).
14. Smewing, G. D., Simonian, K. O. & Gass, I. G. *Contr. Miner. Petrol.* **51**, 49–64 (1975).
15. Beach, A. *J. struct. Geol.* **2**, 427–438 (1980).
16. Jaeger, J. C. & Cook, N. G. W. *Fundamentals of Rock Mechanics*, 247 (Methuen, London, 1969).
17. Cornet, F. H. *Mém. Bur. Rech. Géol. Min.* **91**, 173–182 (1978).

Methane production and simultaneous sulphate reduction in anoxic, salt marsh sediments

Ronald S. Oremland, Lorraine M. Marsh & Sandra Polcin

US Geological Survey, 345 Middlefield Road, Menlo Park, California 94025, USA

It has been generally believed that sulphate reduction precludes methane generation during diagenesis of anoxic sediments^{1,2}. Because most biogenic methane formed in nature is thought to derive either from acetate cleavage or by hydrogen reduction of carbon dioxide^{3–6}, the removal of these compounds by the energetically more efficient sulphate-reducing bacteria can impose a substrate limitation on methanogenic bacteria^{7–9}. However, two known species of methanogens, *Methanosarcina barkeri* and *Methanococcus mazei*, can grow on and produce methane from methanol and methylated amines^{10–13}. In addition, these compounds stimulate methane production by bacterial enrichments from the rumen^{11,14} and aquatic muds^{13,14}. Methanol can enter anaerobic food webs through bacterial degradation of lignins¹⁵ or pectin¹⁶, and methylated amines can be produced either from decomposition of substances like choline, creatine and betaine^{13,14} or by bacterial reduction of trimethylamine oxide¹⁷, a common metabolite and excretory product of marine animals. However, the relative importance of methanol and methylated amines as precursors of methane in sediments has not been previously examined. We now report that methanol and trimethylamine are important substrates for methanogenic bacteria in salt marsh sediments and that these compounds may account for the bulk of methane produced therein. Furthermore, because these compounds do not stimulate sulphate reduction, methanogenesis and sulphate reduction can operate concurrently in sulphate-containing anoxic sediments.

Sediments were taken from the upper 15 cm of a salt marsh located in San Francisco Bay at Palo Alto, California. The black, highly reduced sediments contained methane (~8 ml per l wet sediment; T. Vogel *et al.*, unpublished data) and smelled strongly of hydrogen sulphide. The sediments also contained abundant plant material (such as roots, rhizomes, stems and leaves) of *Spartina foliosa*. Glass jars were filled completely with sediment (to exclude air) sealed and transported back to the laboratory. Experiments began within 1 h of sample collection. Sediments were homogenized anaerobically¹⁸ with San Francisco Bay water (salinity = 20‰; sulphate = 18 mM) to pro-

Table 1 Effect of methanol and sulphate on formation of methane and sulphide by sediment slurries incubated for 6 days

Conditions	CH ₄ ($\mu\text{mol per flask}$)	S ⁻² ($\mu\text{mol per flask}$)
Methanol + sulphate	192 $\pm$ 28	126 $\pm$ 13
Methanol - sulphate	162 $\pm$ 16	104 $\pm$ 22
No amendments + sulphate	0.90 $\pm$ 0.05	111 $\pm$ 22

Results represent the mean $\pm$ 1 s.d. of three flasks. Sulphide was measured by cadmium analysis of acidified sediments²⁴ and initial levels were $\sim 9 \mu\text{mol S}^{-2}$ per flask. The artificial media contained (g l^{-1}): NaCl, 12; $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 5.5; $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.75; KCl, 0.38; NaBr, 0.04; NaHCO_3 , 0.25; Na_2SO_4 or NaCl, 2.8; trace element solution²⁵, 10 ml. Erlenmeyer flasks (125 ml) contained 50 ml of the media plus 25 ml of the sediment homogenate (1:1; sediment: media)¹⁸. Sulphide produced in the sulphate-free flasks was probably from organic sources and reduction of interstitial sulphate ions that were originally present in the sediment sample. Sulphide production was stimulated by amendment with either acetate or H_2 (S.P. and R.S.O., in preparation). Hydrogen also stimulated CH_4 production by these sediments. After 5 days, H_2 -incubated slurries produced more CH_4 ($8.2 \pm 5.3 \mu\text{mol per flask}$) than N_2 -incubated slurries ($0.098 \pm 0.007 \mu\text{mol per flask}$). However, much more H_2 was consumed by the slurries ($506 \pm 61 \mu\text{mol}$ than required for methanogenesis, because of H_2 oxidation by sulphate-reducing bacteria⁷.

duce a dense homogenate which was dispensed into Erlenmeyer flasks and sealed under N_2 . In some experiments, the homogenate was further diluted with bay water to yield a slurry¹⁸. To test the influence of sulphate on methanogenesis from methanol or trimethylamine, dilute slurries were also prepared by homogenizing sediment in a sulphate-free artificial media and dispensing the homogenate into flasks containing more of the same media either amended with or lacking sulphate (Table 1). Additions were made to selected flasks, where indicated, of homogenized *Spartina* material, ^{14}C -methanol and methanol. 2-bromoethanesulphonic acid (BES) was added to certain flasks to inhibit methanogenic bacteria^{18,19}. BES functions as a specific inhibitor of methanogens because the compound is a structural analogue of coenzyme M (2-mercaptoethanesulphonic acid)¹⁹; coenzyme M occurs only in methanogenic bacteria²⁰. All flasks were incubated in the dark at 19°C with constant rotary shaking (200 r.p.m.).

Incubation of sediment homogenates in the presence of BES inhibited methane formation and caused increases in the pool sizes of methanol and trimethylamine with time (Fig. 1). In contrast, uninhibited flasks produced methane and had lower, fluctuating pool sizes of both methanol and trimethylamine. Hydrogen did not accumulate in appreciable quantities (~ 100 – $200 \text{ nmol per flask}$; day 2) in either the BES or the uninhibited flasks, and the gas was not detected by day 3 in any of the flasks. In a second experiment, addition of ^{14}C -methanol to freshly prepared sediment homogenates resulted in the immediate, linear production of $^{14}\text{CH}_4$ and production of both $^{14}\text{CH}_4$ and CH_4 were inhibited by BES (Fig. 2).

In experiments with dilute sediment slurries, methanol-amended flasks ($500 \mu\text{mol per flask}$) produced methane and production levelled off after 5 days incubation. Levels of methane formed in replicate flasks (310 and $340 \mu\text{mol per flask}$) accounted for an 83–91% conversion of the methanol, based on the reaction:



Flasks incubated with methanol plus BES (3.5 mM) did not produce significant amounts of methane ($<0.5 \mu\text{mol CH}_4$ per flask after 2 weeks incubation). Methane formation (from methanol) by sediment slurries incubated in artificial media was not influenced by the presence of sulphate. The rates and extent of methane formation with time were identical for flasks which either contained or lacked sulphate ions (20 mM) and the levels of methane achieved at the end of the experiment (6 days) were comparable (Table 1). Amendment of these slurries with methanol greatly stimulated methanogenesis but did not stimulate sulphate reduction (as measured by S^{-2} ; Table 1). We have obtained similar results for trimethylamine using slurries pre-

pared from nearby intertidal mudflat sediments (S.P. and R.S.O. in preparation).

The accumulation of trimethylamine and methanol in BES-inhibited homogenates (Fig. 1) indicates that these compounds are usually present in these sediments, but are continuously removed and converted to methane by methanogenic bacteria. In addition, no lag of $^{14}\text{CH}_4$ production occurred on addition of ^{14}C -methanol to sediment homogenates (Fig. 2). Because production of $^{14}\text{CH}_4$ was immediate and linear, the methanogens present in the sediments were substrate adapted to methanol and presumably, trimethylamine (^{14}C -trimethylamine was not tested). Furthermore, because 4 mol of trimethylamine can produce 9 mol of methane¹³, the large pool sizes of trimethylamine (Fig. 1b) at day 2 ($\sim 57 \mu\text{mol per flask}$, corresponding to $\sim 570 \mu\text{mol l}^{-1}$) could account for all the methane produced ($\sim 37 \mu\text{mol CH}_4$ per flask at day 3) with 10% coming from methanol. Hydrogen was not an important energy source for methanogens because there was no appreciable accumulation of this gas either in the presence or absence of BES. Disappearance of hydrogen by day 3 therefore, was due to removal by sulphate reducers and other anaerobic bacteria⁷⁻⁹.

As methanol and trimethylamine stimulate methanogenesis, but not sulphate reduction, the two processes can operate simultaneously and independently within anoxic sediments containing these compounds and sulphate ions. It is significant that other workers have observed a rapid conversion of ^{14}C -methylamine to ^{14}C -methane in sulphate-containing marine sediments (M. Winfrey and D. Ward, personal communication). Thus, in sediment pore waters containing relatively high concentrations of sulphate, sulphate reduction proceeds at the expense of short chain fatty acids and hydrogen⁹ while methanogenesis can occur simultaneously through metabolism of non-competitive substrates, such as methanol and methylated amines. As sulphate becomes depleted from sediment pore waters, short chain fatty acids and hydrogen can enter

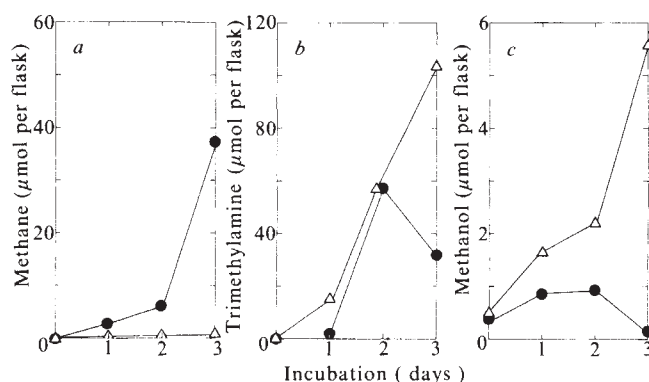


Fig. 1 Levels of methane (a), trimethylamine (b) and methanol (c) in uninhibited flasks (●) and flasks containing 0.7 g BES (Δ). Flasks (total volume, 138 ml) contained 80 ml sediment homogenate (1:1; sediment: bay water), 10 ml bay water and 20 ml homogenized *S. foliosa* materials (1:1 with bay water). A set of eight flasks (four uninhibited; four + BES) were incubated with constant shaking (dark, 19°C). Two flasks were used daily for methanol and trimethylamine determinations. Homogenate was centrifuged (4°C) and supernatant distilled (90– 100°C). Fractions of distillate (100–500 μl) were collected and analysed by flame ionization gas chromatography¹⁸ using 5 μl injections and a $183 \times 0.64 \text{ cm}$ stainless-steel Porapak Q column (110°C). Compounds were identified by retention time. Per cent recoveries of methanol (75%) and trimethylamine (46%) were obtained by distillation of standard solutions, however, distillation efficiencies were not applied to pool size estimates for the purpose of being conservative. The presence of trimethylamine in the distillate was confirmed by mass spectrometry (Accurex Co.). One μl of sample was purged onto a Tekman LSC 2 and thermally absorbed onto 0.5% carboxypak 1500 carboxypak column. The column was ramped from 60 to 200°C and effluent detected on a Finnegan 1020 gas chromatograph-mass spectrometer scanning 41–275 AMU every 2 s. The samples had AMU peaks at 59 (trimethylamine) and were identical with prepared standards. Methanol also accumulated in flasks that contained only homogenized sediments (no added plant materials). After 7 days incubation, 5.5 and 332 nmol CH_3OH were recovered from uninhibited and BES flasks, respectively (trimethylamine was not searched for). The flasks had produced 3,200 nmol CH_4 (uninhibited) and 300 nmol CH_4 (BES) after 1 week.

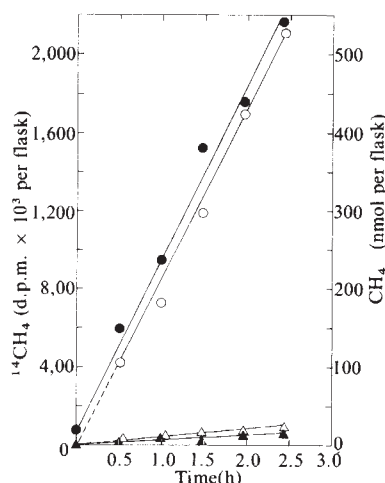


Fig. 2 Formation of CH_4 and $^{14}\text{CH}_4$ by homogenates amended with ^{14}C -methanol (12.5 μCi per flask; 0.28 μmol per flask; ICN Co., initial specific activity 45 mCi mmol^{-1}). Uninhibited flasks: $^{14}\text{CH}_4$, $\bullet$; CH_4 , $\circ$. Flasks inhibited with BES (1 g): $^{14}\text{CH}_4$, $\blacktriangle$; CH_4 , $\triangle$. Flasks (total volume, 263 ml) contained 145 ml of sediment homogenate (1:1 with bay water), 25 ml bay water and 65 ml of homogenized *S. foliosa* materials (1:1 with bay water). Flasks were incubated in the dark (20 $^\circ\text{C}$) with constant rotary shaking (200 r.p.m.). $^{14}\text{CH}_4$ and $^{14}\text{CO}_2$ were determined by gas chromatography–gas proportional counting procedures²⁷ and CH_4 by gas chromatography¹⁸.

methanogenic pathways. Therefore, the ecological niche occupied by methanogens seems broader than previously predicted^{1,2}. As most studies on sediment methanogenesis used ^{14}C -acetate or ^{14}C -bicarbonate as precursors of methane^{6,21}, results can be misleading with regard to zones of methane production and may also underestimate carbon budgets. Future studies should be directed at determining the contribution methanol and methylated amines make to the methane formed in various aquatic sediments. Because carbon isotope enrichment factors are significantly greater in methanol-grown methanogens (~70%)^{23,26} than in hydrogen plus carbon dioxide-grown cells (~40%)²², such information would be of importance for the interpretation of carbon isotopic fractionations of methane during methanogenesis in anoxic sediments.

We thank R. Gunsalus, D. Ward and M. Winfrey for helpful discussions, K. Kvenvolden, R. Smith and W. Balch for discussions and manuscript review, and G. Nicoll for the mass spectrometry work. Mention of brand name products does not constitute an endorsement by the US Geological Survey.

Received 22 October; accepted 18 December 1981.

- Martens, C. S. & Berner, R. A. *Science* **185**, 1167–1169 (1974).
- Claypool, G. E. & Kaplan, I. R. in *Natural Gases in Marine Sediments* (ed. Kaplan, I. R.) 99–140 (Plenum, New York, 1974).
- Hungate, R. E. *Arch. Microbiol.* **59**, 158–165 (1967).
- Jeris, J. S. & McCarty, P. L. *J. Wat. Pollut. Cont. Fedn* **37**, 178–192 (1965).
- Smith, P. H. & Mah, R. A. *Appl. Microbiol.* **14**, 368–371 (1966).
- Cappenberg, T. E. & Prins, R. A. *Ant. van Leeuwenhoek, J. Microbiol. Serol.* **40**, 457–469 (1974).
- Oremland, R. S. & Taylor, B. F. *Geochim. cosmochim. Acta* **42**, 209–214 (1978).
- Abram, J. W. & Nedwell, D. B. *Arch. Microbiol.* **117**, 89–92 (1978).
- Sørensen, J. et al. *Appl. env. Microbiol.* **42**, 5–11 (1981).
- Weimer, P. J. & Zeikus, J. G. *Arch. Microbiol.* **119**, 49–57 (1978).
- Patterson, J. A. & Hespell, R. B. *Curr. Microbiol.* **3**, 79–83 (1979).
- Mah, R. A. *Curr. Microbiol.* **3**, 321–326 (1980).
- Hippe, H. et al. *Proc. natn. Acad. Sci. U.S.A.* **76**, 494–498 (1979).
- Neill, A. R. et al. *Biochem. J.* **170**, 529–535 (1978).
- Donnelly, M. I. & Dagley, S. J. *J. Bact.* **142**, 916–924 (1980).
- Schink, B. & Zeikus, J. G. *Curr. Microbiol.* **4**, 387–389 (1980).
- Ström, A. R. et al. *J. gen. Microbiol.* **112**, 315–320 (1979).
- Oremland, R. S. *Appl. env. Microbiol.* **42**, 122–129 (1981).
- Gunsalus, R. P. et al. *Biochemistry* **17**, 2374–2377 (1978).
- Balch, W. E. & Wolfe, R. S. *J. Bact.* **137**, 256–263 (1979).
- Sansome, F. J. & Martens, C. S. *Science* **211**, 707–709 (1981).
- Games, L. M. et al. *Geochim. cosmochim. Acta* **42**, 1295–1297 (1978).
- Rosenfeld, W. D. & Silverman, S. R. *Science* **130**, 1658–1659 (1977).
- Oremland, R. S. & Silverman, M. P. *Geomicrobiol. J.* **1**, 355–372 (1979).
- Wolin, E. A. et al. *J. Biol. Chem.* **121**, 184–191 (1963).
- Oremland, R. S. et al. *Appl. env. Microbiol.* **43**, 462–468 (1982).
- Cubertson, C. W. et al. *Appl. env. Microbiol.* **41**, 396–403 (1981).

Sulphate and sulphate reduction in early Precambrian oceans

E. M. Cameron

Geological Survey of Canada, Ottawa, Ontario, Canada KIA OE8

Sulphate reduction generally causes isotopic fractionation of sulphur¹. Modern sedimentary sulphide is largely produced by biogenic reduction of sulphate and is typically enriched in ^{32}S (ref. 2). This is balanced by excess ^{34}S in the oceanic sulphate reservoir and evaporites³. High-temperature, inorganic reduction of sulphate may also cause fractionation^{4,5}. Since the work of Ault and Kulp⁶, there has been interest in finding the beginnings of sulphate reduction in the sedimentary record. This is important for several reasons. First, sulphate-respiring bacteria are a milestone of evolution^{7,8}. Second, it established the exogenic sulphur cycle in an essentially modern form. This, with the interconnected oxygen and carbon cycles, regulates the composition of atmosphere and oceans^{9–11}. Third, widespread evidence of sulphate reduction in rocks of a given age and younger indicates that sulphate was established as a major constituent of seawater. In addition to identifying a stage in the evolution of an oxygenated environment¹⁰, this has important metallogenic implications. Schidlowski⁸ has recently concluded that dissimilatory reduction commenced at 2,800–3,100 Myr in an Archaean ocean that had relatively high concentrations of sulphate. I review here the published data and present additional sulphur isotope analyses obtained from the early Precambrian of South Africa. These results indicate that sulphate was a minor component of Archaean and early Proterozoic ocean water, probably $<0.001 \text{ mol l}^{-1}$. The concentration had increased by ~2,350 Myr to levels allowing significant biogenic and inorganic fractionation and the partitioning of $^{32}\text{S}/^{34}\text{S}$ in the exogenic cycle.

Information provided by Precambrian evaporitic sediments is limited. The oldest Proterozoic evaporites for which data are published have maximum ages of only 1,300 Myr. These have $\delta^{34}\text{S}$ values of up to 30‰ (refs 12,13), indicating that sulphate reduction was well established at that time. Stratiform barites found in >3,200 Myr rocks on several continents have $\delta^{34}\text{S}$ of ~0‰ indicating no significant isotopic partitioning in the oceans of that time.

Figure 1 gives the means and standard deviations of $\delta^{34}\text{S}$ for groups of sedimentary sulphide of Archaean (>2,500 Myr) and Aphebian (2,500–1,800 Myr) age. Where geological information permits, the latter have been subdivided into 'Aphebian 1' (>2,300 Myr) and 'Aphebian 2' ($\geq 2,300$ Myr). Data have been excluded where the rocks may have been affected by metasomatism. Care must be taken in interpreting the statistical data, as the sample collections vary greatly in size and in stratigraphical and geographical extent.

Several generalized fields have been outlined in Fig. 1, related to the genesis of the sulphides. Magmatic sulphides have $\delta^{34}\text{S}$ values close to 0‰ and low variance. Groups with more positive $\delta^{34}\text{S}$ values, but also low variance, have been identified as the product of partial reduction of ocean sulphate at high temperature, for example, the Red Sea¹⁴ and in Homestake Mine sediments¹⁵. However, depending on the degree of fractionation⁵ and $\delta^{34}\text{S}$ of ocean sulphate, this field can presumably extend into or across the 'magmatic' field.

Sample groups having distinctly negative $\delta^{34}\text{S}$ values and moderate to high variance are attributed to biogenic sulphate reduction in systems open to exchange with ocean sulphate¹⁶. Groups enriched in ^{34}S and with moderate to high variance have been imputed to closed-system sulphate reduction^{17,18}, that presumably may be either inorganic or biogenic.

All of the Archaean and 'Aphebian 1' groups fall into a restricted range close to 0‰ $\delta^{34}\text{S}$ and low variance, characteristic of magmatic sulphur. The possible exceptions are the 2,750 Myr Woman River and Michipicoten iron formations. The greater variation found in these two groups has been ascribed to biological reduction of low concentrations of sulphate within a restricted basin¹⁹. This interpretation has not been universally accepted²⁰⁻²². As the metals of the iron formations were almost certainly derived from thermal springs²³, fractionation of sulphur at high temperature may have been involved. Indeed, sulphide formed by high-temperature reduction will tend to show the greatest variation in $\delta^{34}\text{S}$ at low sulphate concentrations, that is $\Sigma\text{SO}_4^{2-} \approx \Sigma\text{H}_2\text{S}$, because of the sensitivity of fractionation to solution chemistry near the $\text{SO}_4^{2-}/\text{H}_2\text{S}$ boundary⁵.

While the evidence for bacterial reduction in the Archaean is debatable, there is little doubt that during this time there was no quantitatively important partitioning of $^{32}\text{S}/^{34}\text{S}$ in the exogenic cycle, as shown by the data from five continents summarized in Fig. 1 that have means clustered near 0‰. Also, on the basis of the limited published evidence, this condition persisted into 'Aphebian 1' time.

Partitioning was established at the time of deposition of the Karelian schists²⁴ at ~2,200 Myr. Also, the Outokumpu deposit, a submarine exhalative massive sulphide body within the Karelian, dated at 2,250 Myr (ref. 25), shows a wide range of $\delta^{34}\text{S}$ from -19.2‰ to 5.8‰ (ref. 26). Unfortunately, some of the most comprehensive data on sedimentary sulphides of Aphebian age come from strata whose age is known only within broad limits, that is at the Homestake Mine¹⁵ and the Pine Creek geosyncline²⁷. The enrichment of ^{34}S in many of the early Precambrian units measured to date is intriguing and requires to be balanced by excess ^{32}S elsewhere²⁸.

To provide further data on sulphur isotope fractionation, I have sampled shaly sediments in the Republic of South Africa, representative of the period 3,300-2,000 Myr. This country contains some of the most complete and best exposed early Precambrian successions. Moreover, cratonic conditions were established at a relatively early date here^{29,30}, giving rise to Archaean sediments of different facies from the more typical 'greenstone belt' type. This is important, as the volcanogenic sulphide which dominates the greenstone belts²⁸ may have obscured any biogenic contribution.

Because sulphide is often difficult to separate physically from carbonaceous shale, the sulphur isotopic data were mainly obtained by burning powdered shale in O_2 at 1,300 °C to produce SO_2 . Analyses of pyrite and/or pyrrhotite separated from 20 of these samples give similar $\delta^{34}\text{S}$ values to the total

burn. The organic carbon and the sulphur contents are quite variable for the samples from the Swaziland and Transvaal Supergroups, with values to 5.5% C and 3.2% S for the former and 9.6% C and 4.5% S for the latter. The West Rand Group samples have lower values for carbon, ranging to 0.4% C and 4.2% S.

Deposition of the Transvaal sediments took place mainly in a shallow, marine environment³¹, while the upper part of the Onverwacht was laid down in shallow water, in part possibly evaporitic³². The depositional site of Witwatersrand strata has been enigmatic, either a shallow inland sea, or a lake³³. But Watchorn³⁴ found tidal features in the West Rand Group and both fluvial and shallow water marine facies. Many of the West Rand shales are low grade iron formations, which has been taken as evidence of marine origin³⁰.

The results (Fig. 2) show $\delta^{34}\text{S}$ values close to 0‰ for all samples up to and including the lowermost part of the Malmani Subgroup. This reflects a magmatic source of the sulphide, either directly from volcanic exhalations, or from detrital sulphide that survived weathering in an O_2 -poor atmosphere. Above this, the character of the isotopic distribution is entirely different, with mainly negative numbers. I suggest that the sulphides from the shales above the isotopic transition were derived from reduction of sulphate, with both biogenic and inorganic processes probably being involved.

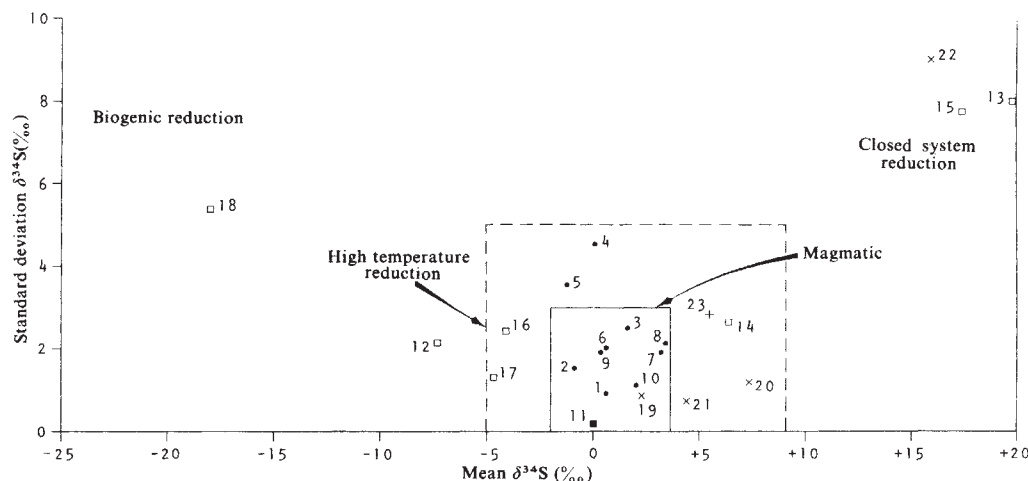
Sulphide from carbonaceous shale of the Timeball Hill Formation is undoubtedly biogenic with whole rock values to -30.6‰ (pyrite from the same sample measured -30.5‰). Two sections were sampled, one being the basal 30 m of the formation in a drill core at Mafefe, northern Transvaal and the other the basal 10 m from a core at Fochville, 400 km distant. The lowest ^{34}S values are from the most sulphidic shales.

Two 20-m sections of pyritic shale from the basal Malips Member of the Penge Iron Formation were taken: one from Mafefe, the other at Penge, 45 km away. They show very similar $\delta^{34}\text{S}$ values with means of -4.8‰ and -4.9‰ respectively and a low standard deviation for all samples of 1.3‰. These absolute values, plus the uniformity of the data, suggest derivation by high-temperature reduction, then thorough mixing before or after exhalation.

The Malmani samples are from carbonaceous shale interbeds in this carbonate unit and are taken from a 1,200-m section near Fochville, plus two samples from Mafefe immediately below the Penge Formation. In part, these data may be interpreted (Fig. 2) as hydrothermal sulphide, similar to the Malips Member. The Malmani is a chemical precursor of the great iron deposits of the Penge and Kuruman Formations, with cycles of iron enrichment in the carbonate³⁵. However, caution

Fig. 1 Sulphur isotope distributions for early Precambrian sedimentary sulphides. ●, Archaean; 1, Isua, Greenland⁶⁹; 2, North Pole, Pilbara Block, and Barberton⁵³; 3, Yilgarn Block²⁰; 4, Michipicoten, Ontario and 5, Woman River, Ontario¹⁹; 6, Rhodesia²¹; 7, Verkhnealdansk and Nimnyr Groups and 8, Fedorovka Group⁴⁸; 9, Onverwacht and Fig Tree Groups; and 10, West Rand Group (see text). ■, 'Aphebian 1': 11, McKim, Frood, and Snider Formations⁷⁰.

□, 'Aphebian 2': 12, Karelian schists²⁴; 13, Krivoi Rog, K3⁷¹; 14, Hammasslahte deposit host rocks⁷²; 15, Onwatin Formation⁷⁰; 16, Malmani Formation (upper 20 samples, Fig. 2); 17 Penge Formation; 18, Timeball Hill Formation (see text). ×, Aphebian undivided: 19, Cahill Formation²⁷; 20 Homestake Formation; 21, Poorman Formation and 22, Ellison Formation¹⁵. +, Modern: 23, Red Sea, hydrothermal (SU1 and SU2 zones)¹⁴. Additional geochronological data from ref. 73.



must be used in interpreting these data, as both upper intertidal and subtidal shales have been identified in the Malmani (ref. 36 and N. J. Beukes personal communication). Each shale type is likely to have a different isotopic distribution. The isotopic transition (Fig. 2) is placed between samples 45 and 77 m above the base of the Malmani. The samples above and below the transition, including those from the Black Reef, are similar, being carbonaceous to highly carbonaceous shales.

The Malmani is bracketed by age determinations of 2,640 Myr for the middle group of the Ventersdorp³⁷ and 2,240 Myr for the Ongeluk lavas of the Pretoria Group (D. Crampton, in ref. 38). Both of these horizons are separated from the Malmani by major unconformities. However, the base of the Malmani is probably closer in age to the Ongeluk than to the Ventersdorp. For this reason the isotopic change is estimated at ~2,350 Myr, a figure likely to be changed with more detailed geochronology.

The isotopic transition at ~2,350 Myr is indicative of one of two possible conditions, the first being an increase in the sulphate concentration of seawater to a level sufficient for reduction to cause significant isotopic fractionation. Harrison and Thode³⁹ showed that the fractionation by *Desulphovibrio desulphuricans* is sharply reduced at $<0.001 \text{ mol}^{-1} \text{ SO}_4^{2-}$, equivalent to ~4% of the concentration in present seawater. For inorganic reduction over the temperature range 200–350 °C at neutral pH, fractionation is minimal when $\Sigma \text{SO}_4^{2-} < \Sigma \text{H}_2\text{S}$ (ref. 5).

The second possible explanation is that sulphate-reducing bacteria evolved at ~2,350 Myr in a pre-existing, sulphate-rich ocean. This is less plausible if it is accepted that the Aphebian also marks the first significant appearance of sedimentary sulphide produced by high-temperature reduction of sulphate, for example, the Homestake Formation¹⁵ and the data given here.

Further evidence of an evolutionary change from a low-sulphate Archaean ocean is the sulphur isotope composition of massive sulphide base metal deposits. Those of Proterozoic and younger age have mean $\delta^{34}\text{S}$ values that commonly differ significantly from 0‰ (refs 40,41), a fact which has been ascribed to derivation, in part at least, by reduction of seawater sulphate^{40,42,43}. Late Archaean (~2,700 Myr) massive sulphides have mean values close to 0‰ (ref. 41), although seawater also had a major role in their formation⁴⁴. Sulphate minerals are absent in these Archaean deposits^{45,46}, whereas they are relatively common in younger deposits.

Possibly contrary evidence is the occurrence of stratiform barite in the Archaean of southern Africa^{32,47–51}; in the Pilbara Block of western Australia^{52–54}; in India⁵⁵; and in the USSR⁴⁸. Where reliable age determinations are available, it seems that these sulphates occur in older Archaean strata, that is $>3,200$ Myr.

Studies carried out on the little metamorphosed Pilbara and Barberton occurrences indicate a depositional environment unusual for the Archaean. That is, a rather extensive, tectonically stable shallow-water shelf with volcanic vents but little detritus⁵⁴. At least some sulphate was deposited in evaporitic conditions as gypsum, then replaced by silica and barite^{53,54,56}. Younger Archaean strata, such as that which dominates the Canadian Shield, lack barite⁵⁰ and shallow water environments of the type described above are rare.

The Pilbara and Barberton barites probably were deposited in restricted basins where sulphate produced, for example, by bacterial oxidation of sulphide⁴⁷, could increase to higher concentration than in the ocean. Mean $\delta^{34}\text{S}$ values of 3.6–3.8‰ for these barites⁵³ indicates that isotopic partitioning was not established in the early ocean. The model discussed below linking global tectonics with ocean sulphate content does not, however, preclude short-term increase in sulphate within Archaean oceans in response to any temporary decline in rates of plate formation.

The data presented here indicate that during Archaean time sulphate was a minor constituent of seawater, probably $<0.001 \text{ mol l}^{-1}$, except within restricted basins. This condition

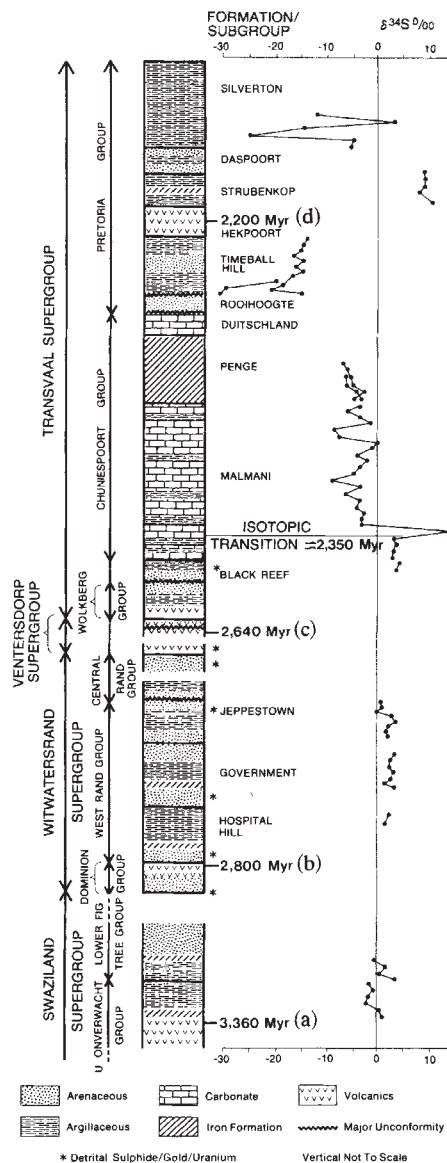


Fig. 2 Sulphur isotope results, early Precambrian argillaceous sediments, Republic of South Africa. Stratigraphical thicknesses not to scale. Isotopic data plotted in true stratigraphical order but not to represent a stratigraphical position within any given unit. Geochronological data: (a) ref. 74; (b) ref. 75; (c) ref. 37; (d) D. Crampton in ref. 38.

persisted into the Proterozoic, but by ~2,350 Myr, sulphate had reached a level sufficient for biogenic or inorganic reduction to cause significant isotopic fractionation and the partitioning of $^{32}\text{S}/^{34}\text{S}$ in the exogenic cycle.

The change in the sulphate concentration of the ocean coincides with a suggested initial increase in the oxygen content of the atmosphere above the primitive level of $<(10^{-2}-10^{-6})$ PAL required for the deposition of placer uraninite⁵⁷. The isotopic transition (Fig. 2) occurs immediately above the youngest occurrence in South Africa of detrital sulphides and uraninite in the Black Reef^{58,59}. In the Huronian of Canada a change from strata hosting pyritic-uraniferous conglomerates to those with black (oxide) sands and red beds occur at ~2,300 Myr (ref. 60).

The transition, during the Lower Proterozoic, from an anoxic atmosphere to one with low levels of O_2 has been attributed to the filling of oceanic sinks for O_2 , notably Fe^{2+} derived from continental weathering^{61–64}. It is curious that this transition and the increase in oceanic sulphate discussed here, approximately

coincided with the greatest tectonic milestone in geological history: the change from a highly mobile Archaean crust to a Proterozoic regime of stable continental masses. Consideration must be given to alternative models, relating atmospheric and hydrospheric evolution to tectonic change.

The dominant mechanism for maintaining O_2 and SO_4^{2-} at low levels during the Archaean may have been a greater rate of exchange of reduced materials between the mantle and the ocean. Edmond *et al.*⁶⁵ estimate the flux of SO_4^{2-} into modern submarine geothermal systems as equivalent to that entering the oceans from rivers. Moreover, major amounts of Fe^{2+} and Mn^{2+} are introduced into the oceans from the ridges.

As the driving force for the exchange between mantle and ocean is heat flow, one would expect this to be greater in the Archaean. Furthermore, there is reason to believe that the exchange was even greater in the Archaean than might be predicted from a linear relationship with the Earth's heat flow. Burke and Kidd⁶⁶ suggest that during this period a greater proportion of the Earth's heat was dissipated at ocean ridges. Bickle⁶⁷ calculated that plate formation was about six times greater at 2,800 Myr than at present, compared with a factor of about three for heat flow. The Archaean/Proterozoic boundary seems to mark a change from a regime of rapid plate formation to a slower one⁶⁸ that would, in turn, produce a lower rate of exchange between reduced mantle material and the hydrosphere/atmosphere. I suggest that it is this change that allowed SO_4^{2-} to become a major component of ocean water and the entry of significant amounts of free oxygen into the atmosphere.

I thank many South African geologists for cooperation and hospitality, in particular, Professor C. R. Anhaeusser, Professor D. A. Pretorius, Dr N. J. Beukes, Professor A. O. Fuller, Dr G. L. Coetzee, P. A. Harrison, Dr W. E. L. Minter, Dr T. G. Molyneux, P. R. Strydom, J. L. Matthysen, P. J. Schoeman, J. P. van Zyl, K. J. Barnard, T. Hopkins, M. Wuth, J. Wigand, P. Becker, L. N. J. Engelbrecht, and Dr P. J. Smit. The isotopic analyses were done by Dr R. Reesman. I thank the Geological Survey of Canada for support, and Professor R. M. Garrels for stimulating discussion on the geochemistry of sulphur.

Received 12 October 1981; accepted 11 January 1982.

- Goldhaber, M. B. & Kaplan, I. R. in *The Sea* Vol. 5 (ed. Goldberg, E. D.) 569–655 (Wiley, New York, 1974).
- Kaplan, I. R., Emery, K. O. & Rittenberg, S. C. *Geochim. cosmochim. Acta* **27**, 297–331 (1963).
- Holser, W. T. & Kaplan, I. R. *Chem. Geol.* **1**, 93–135 (1966).
- Ohmoto, H. *Econ. Geol.* **67**, 551–579 (1972).
- Ohmoto, H. & Rye, R. O. in *Geochemistry of Hydrothermal Ore Deposits* (ed. Barnes, H. L.) 509–567 (Wiley, New York, 1979).
- Ault, W. U. & Kulp, J. L. *Geochim. cosmochim. Acta* **16**, 201–235 (1959).
- Broda, E. *The Evolution of Bioenergetic Processes* (Pergamon, Oxford, 1978).
- Schidlowski, M. *Origins Life* **9**, 299–311 (1979).
- Holland, H. D. *Geochim. cosmochim. Acta* **37**, 2605–2616 (1973).
- Garrels, R. M. & Perry, E. A. in *The Sea* Vol. 5 (ed. Goldberg, E. D.) 303–336 (Wiley, New York, 1974).
- Schidlowski, M., Junge, C. E. & Pietrek, H. J. *Geophys.* **82**, 2557–2565 (1977).
- Whelan, J. F. & Rye, R. O. *Geol. Soc. Am. Abstr. Prog.* **10**, 515 (1978).
- Claypool, G. E., Holser, W. T., Kaplan, I. R., Sakai, H. & Zak, I. *Chem. Geol.* **28**, 199–260 (1980).
- Shanks, W. C. III & Bischoff, J. L. *Econ. Geol.* **75**, 445–459 (1980).
- Rye, D. M. & Rye, R. O. *Econ. Geol.* **69**, 293–317 (1974).
- Schwarz, H. P. & Burnie, S. W. *Miner. Deposita* **8**, 264–277 (1973).
- Thode, H. G. & Monster, J. *Am. Ass. petrol. Geol. Mem.* **4**, 367–377 (1965).
- Burnie, S. W., Schwarz, H. P. & Crockett, J. H. *Econ. Geol.* **67**, 895–914 (1972).
- Goodwin, A. M., Monster, J. & Thode, H. G. *Econ. Geol.* **71**, 870–891 (1976).
- Donnelly, T. H. *et al. J. geol. Soc. Aust.* **24**, 409–420 (1977).
- Fripp, R. E. P., Donnelly, T. H. & Lambert, I. B. *Geol. Soc. S. Afr. Spec. Publ.* **6**, 205–208 (1979).
- Ashendorf, D. *Origins Life* **10**, 325–333 (1980).
- Goodwin, A. M. *Econ. Geol.* **68**, 915–933 (1973).
- Kouvo, O. *Comm. Geol. Finland Bull.* **182**, 1–70 (1958).
- Kouvo, O. & Kulp, J. L. *Ann. N. Y. Acad. Sci.* **91**, 476–491 (1961).
- Makela, M. *Comm. Geol. Finland Bull.* **267**, 1–45 (1974).
- Donnelly, T. H. & Ferguson, J. in *Uranium in the Pine Creek Geosyncline* (eds Ferguson, J. J. & Goleby, A. B.) 397–406 (IAEA, Vienna, 1980).
- Cameron, E. M. & Garrels, R. M. *Chem. Geol.* **28**, 181–197 (1980).
- Pretorius, D. A. *Econ. Geol. Res. Unit Inf. Circ.* **87**, 1–22 (Univ. Witwatersrand, 1974).
- Cloud, P. *Geol. Soc. S. Afr. Trans. Annex* **79**, 1–32 (1976).
- Button, A. *Miner. Sci. Engng* **8**, 262–293 (1976).
- Lowe, D. R. & Knauth, L. P. *J. Geol.* **85**, 699–723 (1977).
- Pretorius, D. A. *Geol. Soc. S. Afr. Spec. Publ.* **6**, 33–55 (1979).
- Watchorn, M. B. *Econ. Geol. Research Unit Inf. Circ.* **148**, 9 (University Witwatersrand, 1980).
- Button, A. *Econ. Geol.* **71**, 193–201 (1976b).
- Buekes, N. J. *Sedim. Geol.* **18**, 201–221 (1977).

- Van Niekerk, C. B. & Burger, A. J. *Geol. Soc. S. Afr. Trans.* **81**, 155–163 (1978).
- Hunter, D. R. *Precamb. Res.* **1**, 295–326 (1974).
- Harrison, A. G. & Thode, H. G. *Trans. Faraday Soc.* **53**, 1648–1651 (1957).
- Sangster, D. F. in *Handbook of Stratabound and Stratiform Deposits* Vol. 2 (ed. Wolfe, K. H.) 219–266 (Elsevier, Amsterdam, 1976).
- Sangster, D. F. *Geol. Ass. Can. Spec. Pap.* **20**, 723–739 (1980).
- Sasaki, A. *Geochem. J.* **4**, 41–51 (1970).
- Sasaki, A. & Kajiwara, Y. *Soc. Mining Geol. Jap. Spec. Iss.* **3**, 289–294 (1971).
- Large, R. R. *Econ. Geol.* **72**, 549–572 (1977).
- Sangster, D. F. *Geol. Surv. Can. Pap.* **72**, 22–1–44 (1972).
- Hutchison, R. W. *Econ. Geol.* **68**, 1223–1246 (1973).
- Perry, E. C., Monster, J. & Reimer, T. *Science* **171**, 1015–1016 (1971).
- Vinogradov, V. I., Reimer, T. O., Leites, A. M. & Smelov, S. B. *Lithol. Miner. Resour.* **11**, 407–420 (1976).
- Heinrichs, T. K. & Reimer, T. O. *Econ. Geol.* **72**, 1426–1441 (1977).
- Thorpe, R. I. *Econ. Geol.* **74**, 700–702 (1979).
- Reimer, T. O. *Precamb. Res.* **12**, 393–410 (1980).
- Hickman, A. H. *Geol. Surv. Ann. Rep.* 1972, 57–60 (1973).
- Lambert, I. B., Donnelly, T. H., Dunlop, J. S. R. & Groves, D. I. *Nature* **276**, 808–810 (1978).
- Barley, M. E., Dunlop, J. S. R., Glover, J. E. & Groves, D. I. *Earth planet. Sci. Lett.* **43**, 74–84 (1979).
- Radhakrishna, B. P. & Vasudev, V. N. *J. geol. Soc. Ind.* **18**, 525–541 (1977).
- Dunlop, J. S. R. *Publ. Geol. Dept. Extens. Ser. Univ. West. Aust.* **2**, 1–30 (1978).
- Grandstaff, D. E. *Precamb. Res.* **13**, 1–26 (1980).
- Papenfus, J. A. in *The Geology of Some Ore Deposits in Southern Africa* Vol. 1 (ed. Haughton, S. H.) 191–218 (Geological Society of South Africa, 1964).
- de Waal, S. A. & Herzberg, W. *Geol. Surv. S. Africa Ann.* **7**, 111–124 (1968).
- Roscoe, S. M. *Geol. Ass. Can. Spec. Pap.* **12**, 31–47 (1973).
- Cloud, P. E. *Science* **160**, 729–736 (1968).
- Cloud, P. E. *Econ. Geol.* **68**, 1135–1143 (1973).
- Schidlowski, M., Eichmann, R. & Junge, C. E. *Precamb. Res.* **2**, 1–69 (1975).
- Schidlowski, M. in *Origin of Life* (ed. Nola, H.) 3–20 (Centre Acad. Publ., Japan, 1978).
- Edmond, J. M. *et al. Earth planet. Sci. Lett.* **46**, 1–18 (1979).
- Burke, K. & Kidd, W. S. F. *Nature* **272**, 240–241 (1978).
- Bickle, M. J. *Earth planet. Sci. Lett.* **40**, 301–315 (1978).
- Dewey, J. F. & Windley, B. F. *Phil. Trans. R. Soc. A* **301**, 189–206 (1981).
- Monster, J. *et al. Geochim. cosmochim. Acta* **43**, 405–413 (1979).
- Thode, H. G., Dunford, H. B. & Shima, M. *Econ. Geol.* **57**, 565–578 (1962).
- Tugarinov, A. I. & Grinenko, V. A. in *Problems of Geochemistry* (ed. Khitarov, N. I.) 205–216 (Nauka, Moscow 1965).
- Hyvarinen, L., Kinnunen, K. & Makela, M. *Comm. Geol. Finland Bull.* **293**, 1–23 (1977).
- Douglas, R. J. W. *Geol. Surv. Can. Pap.* **80**, 24, 1–19 (1980).
- Van Niekerk, C. B. & Burger, A. J. *Geol. Soc. Afr. Trans.* **72**, 9–21 (1969).
- Van Niekerk, C. B. & Burger, A. J. *Geol. Soc. S. Afr. Trans.* **72**, 37–45 (1969).

Pheromones in mice: reciprocal interaction between the nose and brain

E. B. Keverne & C. de la Riva

Department of Anatomy, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK

Blockade of pregnancy by odours from strange males^{1–3} has interested considerably those that adopt a sociobiological approach to reproduction^{4–6}. It has been suggested that the mechanism has evolved to promote heterogeneity in the population, and that strange males in possessing the capacity to block pregnancy thereby increase their reproductive potential. However, such knowledge as we have of the territorial behaviour⁷ and social organization of mice⁸ makes this explanation less likely as resident males have such an advantage over intruders that access by strange males is probably an infrequent event. Another explanation relates pregnancy block to the effect that the male pheromones have on the female reproductive hormones in other contexts. Male pheromones can stimulate both early puberty⁹ and induction of oestrus^{10–12} in grouped females by suppressing prolactin secretion. Such a response, highly appropriate in this context, would be extremely disadvantageous following fertilization, since lowering prolactin is known to prevent implantation^{13,14}. Thus, some mechanism must exist to offset the more general effect of the male's own pheromone on the endocrine function of his female at such times, and this, we suggest, is prevented by the noradrenergic mechanism which we describe here.

In the first experiment, female mice (BALB/c), which had been housed alone for part of the experiment, were taken from their home cage and placed with BALB/c stud males for 20 min to allow mating. They were then removed to cages containing F₁ male bedding while the stud male was returned to the females' home cage. The F₁ males were derived from crossing C57 females with CBA males, and produce pheromones that

Table 1 Effect of replacing stud male odours immediately after mating on stud males' ability to block pregnancy

Stud male	Exposure to soiled bedding (0–48 h)	Exposure to male (48–72 h)	% Showing pregnancy block (n)	
BALB/c	F ₁	BALB/c	75 (8)	$P < 0.01$ — NS*
BALB/c	BALB/c	BALB/c	0 (8)	
BALB/c	BALB/c	F ₁	80 (10)	
BALB/c	F ₁	—	0 (8)	
BALB/c	F ₁	F ₁	0 (8)	

* The data were subjected to χ^2 test of significance. NS, not significant.

are sufficiently different from those of BALB/c males to prevent pregnancy occurring in BALB/c females. After 48 h the females were returned to their home cages containing the stud males. The control group consisted of females that were treated in the same way, except that, after mating, they were moved to cages containing BALB/c male bedding. Vaginal smears were taken daily to determine the onset of oestrus and 6 days after mating, females were killed and their uteri examined for implantation sites to demonstrate whether an olfactory block to pregnancy had occurred. The control females, which had been returned to the stud male after previous exposure to bedding from the same strain (BALB/c) showed implantation, as expected (Table 1). However, in those females exposed to F₁ male bedding for 48 h after mating, the stud male appeared to have been transformed to a 'strange' one, for subsequent exposure of the females to the stud male prevented implantation of his own fertilized ova.

Neither exposure to F₁ bedding by itself for 48 h immediately after BALB/c mating nor a return to F₁ males after this procedure had any blocking effect on pregnancy. This experiment demonstrates that the odour cues which follow mating are the ones 'recognized' by the female's neuroendocrine system, and that this recognition is necessary for pregnancy block to occur. Stud males can thus cause fetal resorption, providing the female has received prolonged exposure to the odour of a strange male, replacing that of the stud.

How does the act of mating and exposure to odour immediately afterwards 'print' recognition of the stud male, thereby preventing him from blocking his own pregnancy? It is known that the so-called primer effects of such pheromones are transmitted by the vomeronasal accessory part of the olfactory system^{9,12,14–16}, which projects to the cortico-medial amygdala¹⁷ part of the brain closely related to the ventromedial hypothalamus, itself concerned with neuroendocrine activity. Both the accessory and main olfactory bulbs receive a noradrenergic projection¹⁸, and this system has been implicated in certain forms of recognition, for example the modulation of neural activity after exposure to certain types of visual stimuli early in life¹⁹, and the neuroendocrine response of the female rat to tactile stimuli inducing lordosis and pseudo-pregnancy²⁰.

To determine whether the centrifugal noradrenergic projection to the olfactory bulbs was involved in the response of the female mice to the males' pheromones, discrete lesions were made by stereotactically injecting the neurotoxin 6-hydroxydopamine (2 μ g 6-OHDA in 0.5 μ l saline containing 2 μ g ml⁻¹ ascorbic acid for 5 min) bilaterally into the medial olfactory

stria which carry the noradrenergic input to the olfactory and accessory olfactory bulbs. A second group of mice were injected in the same way with vehicle alone, a third group with neurotoxin directly into the accessory olfactory bulb, and a fourth group received no injections. After at least 4 days, females from each group were housed in pairs with a BALB/c stud male, which was removed the morning after mating had occurred (as assessed by observing vaginal plugs in the female); 24 h later the stud male was returned to the fertilized female and housed with her for 48 h. A fifth control group received 6-OHDA lesions to the medial olfactory stria but were kept away from all males after mating to determine whether the procedure itself interfered with pregnancy. Vaginal smears were taken daily, all females were killed 6 days after mating and the uterus examined for implantation sites; the olfactory and accessory olfactory bulbs were removed and their noradrenaline content measured.

Both groups 1 and 2 (female mice in which noradrenaline had been depleted in the accessory olfactory bulbs) failed to 'recognize' the stud male, and thus pregnancy was blocked by exposure to him, whereas the neurotoxin lesion itself (group 5) did not interfere with pregnancy (Table 2). As expected, in mice undergoing either sham operations or no treatment (groups 3 and 4), the stud male failed to block pregnancy. Measurement of the noradrenaline content showed that 6-OHDA lesions of the medial olfactory stria produced a significant decrease in noradrenaline (70%) in the olfactory and the accessory olfactory bulbs, without significantly altering dopamine (Table 2). Direct measurement of noradrenaline content in the accessory olfactory bulb was not possible because of its small size.

These experiments show that, in contrast to what is generally believed, a special neuromechanism exists not so much to ensure that strange males will block pregnancy in female mice, but so that familiar (that is, stud) males will not do so. This implies that some neural protective mechanism must be set in motion during mating which ensures that continued presence of the stud male or his odour after coitus will not prevent subsequent implantation. Such a process is dependent on olfactory input to the female which outlasts the act of coitus itself, as we have shown that it can be overridden by a second olfactory stimulus applied immediately after mating has occurred. For the neural mechanism to be effective, there must exist an intact noradrenergic input to the accessory olfactory bulb. The action of this noradrenergic mechanism in the olfactory system may be similar to that of other sensory systems, serving to enhance recognition of the stud male by increasing the signal-to-noise ratio^{21,22}. Alternatively, it has been suggested that noradrenergic mechanisms may have an important role in the memory of events which have survival value²³, a 'print-now' mechanism for incoming sensory information to be stored for later retrieval¹⁹.

The response to male pheromones in female mice has selective advantages for accelerating puberty and induction of oestrus^{7,8}, but carries a risk for pregnancy should contact with these pheromones occur during a 3-day period after mating. The noradrenergic mechanism apparently sets in motion a relatively short-term neuroendocrine memory, which protects the female from pregnancy block which would otherwise be induced by her stud male.

Table 2 Effect of olfactory noradrenergic lesions in the female mouse on the stud males' ability to block pregnancy

Group	Treatment	% Showing pregnancy block (n)
1	6-OHDA (medial olfactory stria)	88 (8)
2	6-OHDA (accessory olfactory bulb)	100 (8)
3	Sham lesion	12 (8)
4	No treatment	0 (8)
5	6-OHDA (medial olfactory stria)	0 (8)

Re-exposure to stud male

No male exposure

NS

NS

NS

NS

NS

NS

NS

NS

NS

NS

Mean noradrenaline content of olfactory bulbs in groups 3 and 1 was 0.37 ± 0.08 and 0.11 ± 0.07 ng per mg tissue, respectively, $t = 6.59$; $P < 0.0004$. Mean noradrenaline depletion = 70%. Dopamine content of olfactory bulbs in groups 1 and 3 was 0.58 ± 0.2 and 0.38 ± 0.06 ng per mg tissue, respectively ($t = 0.82$, not significant).

* The data were subjected to χ^2 tests of significance.

We thank Anne Lloyd-Thomas, Barry Everitt, Gerald Moore and Joe Herbert for thought-provoking discussion.

Received 9 November 1981; accepted 29 January 1982.

1. Bruce, H. M. *Nature* **184**, 105 (1959).
2. Bruce, H. M. *J. Reprod. Fert.* **1**, 96–103 (1960).
3. Bruce, H. M. *J. Reprod. Fert.* **2**, 134–142 (1961).
4. Wilson, E. O. in *Sociobiology, The New Synthesis*, 321 (Harvard University Press, 1976).
5. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) (Aldine, Chicago, 1972).
6. Labov, J. B. *Am. Nat.* **118**, 361–371 (1981).
7. Bronson, F. H. *Q. Rev. Biol.* **54**, 265–299 (1979).
8. Massey, A. & Vandenbergh, J. G. *Science* **209**, 821–822 (1980).
9. Lomas, D. & Keverne, E. B. *J. Reprod. Fert.* (in the press).
10. Bronson, E. H. *Biol. Reprod.* **15**, 147–152 (1976).
11. Charlton, H. M., Milligan, S. R. & Versi, E. *J. Reprod. Fert.* **52**, 283–288 (1978).
12. Reynolds, J. M. & Keverne, E. B. *J. Reprod. Fert.* **57**, 31–35 (1979).
13. Bruce, H. M. *J. Reprod. Fert.* **10**, 141–143 (1965).
14. Bellringer, J. F., Pratt, H. E. & Keverne, E. B. *J. Reprod. Fert.* **59**, 223–228 (1980).
15. Johns, M. A., Feder, H. H., Komisaruk, B. R. & Mayer, A. D. *Nature* **271**, 446–448 (1978).
16. Kaneko, N., Debski, E. A., Wilson, M. C. & Whitten, W. K. *Biol. Reprod.* **22**, 873–878 (1980).
17. Scalia, F. & Winans, S. S. *J. comp. Neurol.* **161**, 31–56 (1975).
18. Fallon, J. H. & Moore, R. Y. *J. comp. Neurol.* **180**, 533–544 (1978).
19. Kasamatsu, T. & Pettigrew, J. D. *Science* **194**, 206–209 (1976).
20. Hansen, S., Stanfield, E. J. & Everitt, B. J. *Nature* **286**, 152–154 (1980).
21. Foote, S. L., Freedman, R. & Oliver, A. R. *Brain Res.* **86**, 229–242 (1975).
22. Waterhouse, B. D. & Woodward, D. J. *Expl Neurol.* **67**, 11–34 (1980).
23. Kety, S. S. in *The Neurosciences, Second Study Program* (ed. Schmidt, F. O.) Rockefeller University Press, New York, 1970).

Extensive elongation of axons from rat brain into peripheral nerve grafts

Martin Benfey & Albert J. Aguayo

Neurosciences Unit, The Montreal General Hospital and Department of Neurology, McGill University, 1650 Cedar Avenue, Montreal, Canada H3G 1A4

The failure of axons to elongate in the injured central nervous system (CNS) of adult mammals restricts drastically the establishment of connections with target tissues situated more than a few millimetres away. Mechanisms that include a primary inability of some nerve cells to support renewed axonal growth, a premature formation of synapses on nearby neurones¹, an obstruction caused by the formation of a glial scar^{2,3}, and other influences of the microenvironment^{4–7} are presumed to contribute to the failure of nerve fibres to regenerate as effectively in the CNS as in the peripheral nervous system (PNS). Support for the hypothesis that conditions in the glial environment of injured fibres have a decisive role in successful axonal elongation has recently come from studies using transplants containing either central glia or peripheral nerve segments as conduits of axon growth^{7,8}. While CNS glial grafts have been shown to prevent growth of PNS fibres^{7–9}, experiments which used labelling techniques to trace the source of axons growing into PNS grafts provided evidence that processes from nerve cells in the spinal cord and medulla oblongata of adult rats may increase in length by 1 or more centimetres when the CNS glial environment is replaced by that of peripheral nerves^{10,11}. Here we report for the first time the extensive elongation of axons from neurones in the brain of adult rats through PNS grafts introduced into the cerebral hemispheres.

In 20 Sprague–Dawley rats weighing ~300 g the end of a 15-mm segment of an autologous sciatic nerve was inserted close to the basal ganglia or to the cerebral cortex (Fig. 1A). The animals had previously been anaesthetized with sodium pentobarbital administered intraperitoneally (5 mg per 100 g body weight). A glass rod held in a micromanipulator was inserted through an opening in the frontal bone and dura to a depth calculated to produce lesions extending to the sensory cortex or basal ganglia. The sciatic nerve graft was then introduced into each of these lesions and its epineurium anchored to the dura mater with 10–0 nylon sutures. The space between the nerve and the surrounding skull was sealed with bone wax and the free outer end of the nerve sutured to the temporalis muscle. The animals survived without any apparent neurological deficit other

than limb weakness resulting from the removal of one sciatic nerve. The rats were examined 5–23 weeks after grafting to determine whether axons from neurones within the brain had elongated along the graft. For this purpose the graft was dissected along its extracranial course, transected 2 mm from its insertion into the temporalis muscle, and the free end placed for 50 min in a pad of Gelfoam soaked in a 20% solution of horseradish peroxidase (HRP; type VI, Sigma). The remaining portion of the graft was placed over a plastic sheath and covered with vaseline to avoid contamination of the surrounding tissue by HRP (Fig. 1A). In five additional animals, used as controls, the mid-portion of the regenerated graft was crushed three times with fine-tip forceps cooled in liquid nitrogen, 9–16 weeks after grafting and half an hour before the application of HRP to the distal end of the crushed nerve. About 48 h after applying HRP, the animals were perfused through the heart with a 0.1 M phosphate buffer solution followed by 3% glutaraldehyde. The brain was removed and rinsed in 10% sucrose buffer overnight. Then, 40 µm-thick coronal sections were cut in a cryostat, reacted with tetramethylbenzidine and hydrogen peroxide¹², counterstained with neutral red and mounted in Permount.

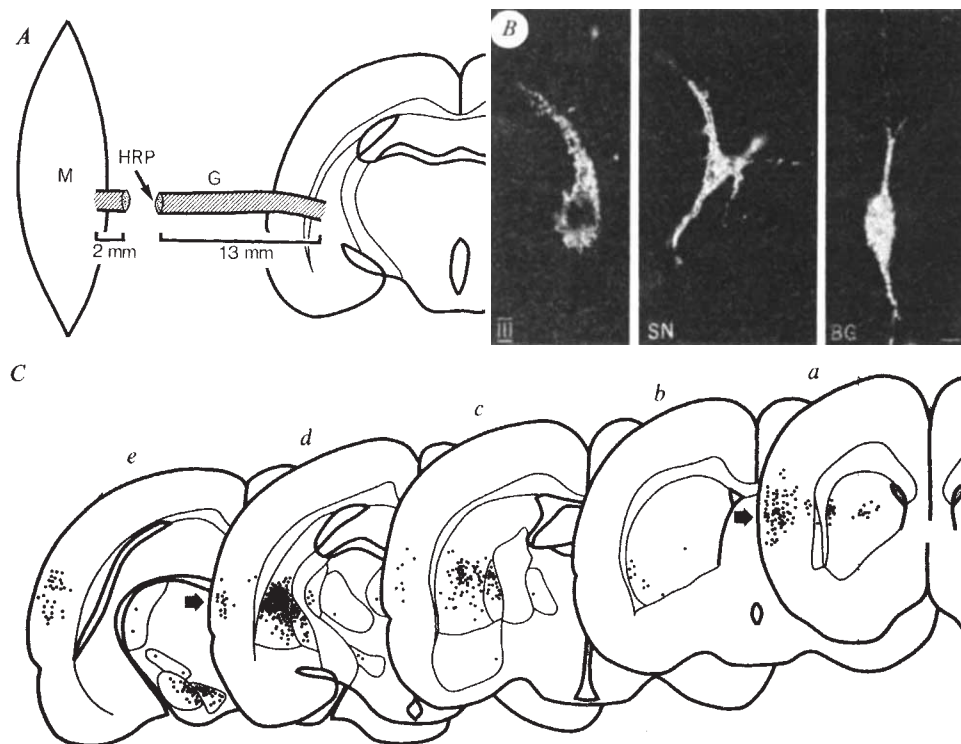
Cross-sections of the outermost portion of each graft were examined using a light microscope and found to contain regenerated fibres surrounded by Schwann cells and myelin. The overall appearance of these fibres resembled that of any regenerated peripheral nerve. HRP-labelled neurones were sought by dark- and brightfield light microscopy in sections rostral to and including the midbrain. An average of 145 sections of the brain were examined per rat. In 10 experimental animals the graft terminated close to the cortex; in 8 it was found within the caudate putamen and in 1 rat each it reached the hippocampus or the ventral nucleus of the thalamus. Of a total of 444 HRP-labelled neurones (Fig. 1B, C), most were found in the caudate putamen and sensory cortex (Table 1). The labelled nerve cells in the sensory cortex were scattered through layers II–VI. Both large and small cortical neurones were labelled but the HRP-containing cells tended to be of larger size than the unlabelled ones in the same regions (Fig. 2). Most (79%) of the labelled cells found in the brain of these animals were within 1.5 mm of the intracerebral tip of the graft; the rest were located between 1.5 and 6.5 mm from the graft. The greatest distance between the graft and the soma of the labelled cells was found for neurones in the substantia nigra (SN). Because all labelled neurones in the experimental rats were confined to a well circumscribed area close to the graft and as only two neurones were labelled in the five controls in which the graft was crushed before HRP application, we believe that most of the cells in the experimental animals used in this study were not falsely labelled by haematogenous spread or by diffusion along the grafted nerve. Thus, on the basis of these HRP tracing studies, we conclude that axons of cells in different regions of the rat brains examined have become elongated by at least 13 mm, the length of nerve graft used in the labelling experiments.

The fact that the largest number of labelled neurones was found in the caudate putamen may reflect a greater ability of

Table 1 Distribution of HRP-labelled cortical and subcortical neurones in the brain of rats having PNS grafts

Site	No. of HRP-labelled neurones
Caudate putamen	218
Globus pallidus	42
Thalamus	10
Substantia nigra	22
Ventral tegmental area	20
Zona incerta	4
Amygdala	1
Clastrum	1
Cortex	126
Total	444

Fig. 1 A, Schematic representation of the experimental design: a PNS graft (G) has been introduced into the rat brain. The intracranial tip of the graft can be positioned close to the cortex or subcortical nuclei within the hemisphere. The outer end of the graft, initially attached to the temporalis muscle (M), is cut to apply HRP. The advantages of this design are: (1) as most of the graft is extracranial, the cells from which the axons growing into the graft have originated can be determined by retrograde labelling with little risk of contamination from spread of HRP into the brain; (2) the length of axons regenerating within the graft can be measured by determining the distance between the site of HRP application and the soma of the labelled cells; (3) the grafts may be positioned in several regions of the brain to test the regenerative capacity of different neuronal populations. B, dark-field light micrographs of HRP-labelled neurones in layer III of the sensory cortex, from the pars compacta of the substantia nigra (SN) and from basal ganglia (BG). Calibration bar, 10 μm . C, transverse sections of the brain, obtained at 1.5-mm intervals, are arranged in a rostral (a) to caudal (e) direction to illustrate the approximate position of the 444 labelled neurones (●) observed in 20 grafted animals. The arrow in d indicates the site of 15 graft insertions made stereotactically 4.4 mm rostral to the inter-aural line in the horizontal plane¹⁹. In five other animals the graft was inserted 9 mm rostral to and 2 mm above these two reference lines (arrow in a).



Calibration bar, 10 μm . C, transverse sections of the brain, obtained at 1.5-mm intervals, are arranged in a rostral (a) to caudal (e) direction to illustrate the approximate position of the 444 labelled neurones (●) observed in 20 grafted animals. The arrow in d indicates the site of 15 graft insertions made stereotactically 4.4 mm rostral to the inter-aural line in the horizontal plane¹⁹. In five other animals the graft was inserted 9 mm rostral to and 2 mm above these two reference lines (arrow in a).

these cells to regenerate in the experimental conditions used, or may simply be the consequence of their proximity to the transplanted nerve; different graft locations may elucidate this question. Most of the labelled cerebral neurones were situated near the tip of the graft, which suggests that the remarkable elongation of their axons results from local sprouting triggered by injury¹³⁻¹⁵ resulting from the grafting procedure and facilitated by growth-favouring conditions in the PNS transplant. It is also possible that growth factors released by the segment of nerve^{16,17} may influence only the soma of nearby cells. The demonstration that the cerebral neurones that give rise to the re-growing axons are confined to an area within a few millimetres of the end of the graft is consistent with similar observations in the rat spinal cord and brain stem^{10,11}. These findings suggest that it may be possible to use similar grafting methods to investigate further neuronal plasticity in other specific areas of the CNS of mature and immature animals and in conditions associated with ageing and disease.

Although the axons of neurones labelled in these experiments would, in the intact animal, reside and project exclusively within the CNS, they have successfully elongated along the implanted grafts in a manner similar to that of fibres injured in peripheral nerves. This indicates that such elongation may depend on properties shared by populations of intrinsically and extrinsically projecting nerve cells, the potential for the growth of intrinsic CNS neurones being expressed when the neuroglial environment in the brain is substituted by that in peripheral nerves. It is not known, however, whether the cells labelled in these experiments indicate a more general potential for regeneration by CNS neurones when interacting with Schwann cells and other PNS components.

These results also suggest that in conditions created by nerve grafting, the axons of some CNS neurones can grow to lengths greater than those normally reached in the intact animal. Axons from the striatal neurones, which are thought normally to project only to the nearby globus pallidus and SN¹⁸, appear to have grown along the entire length of the grafts. Furthermore, because the intracerebral tip of the 13-mm-long graft was as far as 6.5 mm from the labelled SN cells in some animals, we

calculate that the axons from nigral neurones must have attained a total length of almost 20 mm. The extent of this axonal growth can be placed in perspective by comparing it with the maximum direct distance between the SN and the cerebral cortex: 14–15 mm in 300 g rats¹⁹.

The hypothesis that the microenvironment at the tip of the

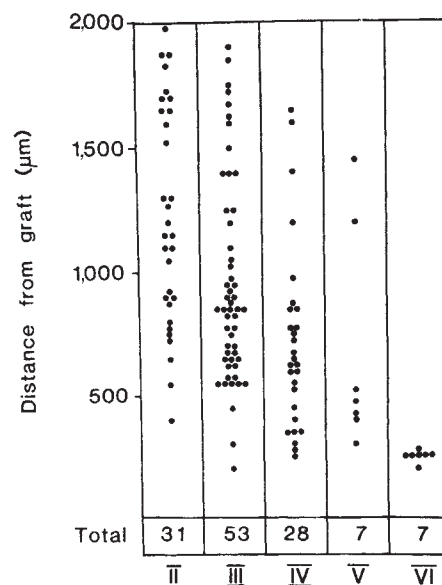


Fig. 2 Distribution of 126 HRP-labelled neurones in layers II–VI of the somatosensory cortex of grafted rats. The vertical axis indicates the distance between the labelled soma and the intracerebral tip of the graft. The cells in the more superficial layers are nearly 2 mm from the graft. The diameter of the labelled cortical neurones averaged $24.9 \pm 4.05 \mu\text{m}$ (mean $\pm$ s.d.) (range 16.8–38.4 μm) while that of 250 unlabelled neighbouring cells was $17.9 \pm 3.63 \mu\text{m}$ (range 12.0–28.8 μm). The difference in size between the labelled and unlabelled cells may result from a neuronal reaction to axon interruption at the time of HRP application to the graft.

re-growing axon has a decisive role in the success or failure of regeneration implies that non-neuronal PNS components influence neuronal mechanisms that regulate elongation^{20,21} whereas the CNS milieu of adult mammals either lacks these growth-promoting factors or exerts an inhibitory influence. The critical differences between the CNS and PNS microenvironment are unknown but may include specific substances released by cells^{6,22}, influences arising from extracellular components²³, surface properties of sheath cells²⁴ and, possibly the different spatial arrangement of glia in the CNS and PNS. Further questions which require investigation are whether central axons elongate into PNS grafts by the regrowth of the damaged fibres, by collateral sprouting or by both. Will long projecting axons

such as those of the corticospinal tract, also grow along PNS grafts after they are interrupted in the brain? What connections, if any, can be established between the growing CNS axons and the target tissues to which the nerve graft is attached? Our findings suggest that the elongation of axons from these central neurones is governed by interactions between the growing nerve fibres and their surrounding tissues, and that abortive regeneration is not the result of an intrinsic lack of neuronal potential to support renewed growth.

This work was supported by grants from the MRC of Canada, the Muscular Dystrophy Association of Canada and the Multiple Sclerosis Society of Canada. M.B. held a studentship from the Faculty of Medicine of McGill University.

Received 11 November 1981; accepted 14 January 1982.

- Bernstein, J. J. & Bernstein, M. E. *Expl Neurol.* **30**, 336–361 (1971).
- Windle, W. F. *Physiol. Rev.* **36**, 426–440 (1956).
- Clemente, C. *Int. Rev. Neurobiol.* **6**, 257–301 (1964).
- Cajal, S. R. *Degeneration and Regeneration of the Nervous System* (ed. May, R. M.) (Oxford University Press, London 1928).
- Tello, F. *Trab. Lab. Invest. biol. Univ. Madr.* **9**, 123–159 (1911).
- Varon, S. *Expl Neurol.* **54**, 1–6 (1977).
- Aguayo, A. J., Bray, G. M., Perkins, C. S. & Duncan, I. D. *Soc. Neurosci. Symp.* **4**, 361–383 (1979).
- Aguayo, A. J., David, S., Richardson, P. & Bray, G. M. *Advances in Cellular Neurobiology* (eds Fedoroff, S. & Hertz, L.) Vol. 3, 215–234 (Academic, New York, 1982).
- Weinberg, E. L. & Spencer, P. S. *Brain Res.* **162**, 273–279 (1979).
- Richardson, P. M., McGuinness, U. M. & Aguayo, A. J. *Nature* **284**, 264–265 (1980).
- David, S. & Aguayo, A. J. *Science* **214**, 931–933 (1981).
- Mesulam, M.-M. *J. Histochem. Cytochem.* **26**, 106–117 (1978).
- Liu, C. N. & Chambers, W. W. *Archs Neurol. Psychiat.*, *Lond.* **79**, 46–61 (1958).
- Raisman, G. & Field, P. M. *Brain Res.* **50**, 241–264 (1973).
- Goldberger, M. E. & Murray, M. in *Neuronal Plasticity* (ed. Cotman, C. W.) 73–96 (Raven, New York, 1978).
- Lundborg, G., Longo, F. L. & Varon, S. *Brain Res.* **232**, 157–161 (1982).
- Ebendal, T. & Richardson, P. M. *Proc. 11th ann. Meet. Soc. Neurosci.* (1981).
- Graybiel, A. M. & Ragsdale, C. W. *Prog. Brain Res.* **51**, 239–284 (1979).
- Pellegrino, L. J., Pellegrino, A. S. & Cushman, A. J. *A Stereotaxic Atlas of the Rat Brain* (Plenum, New York, 1979).
- Lasek, R. J. & Hoffman, P. N. *Cell Motility* (eds Goldman, R., Pollard, T. & Rosenbaum, J.) 1021–1051 (Cold Spring Harbor Laboratory, New York, 1976).
- Grafstein, B. & McQuarrie, I. G. in *Neuronal Plasticity* (ed. Cotman, C. W.) 155–195 (Raven, New York, 1978).
- Varon, S. S. & Bunge, R. P. *A. Rev. Neurosci.* **1**, 327–361 (1978).
- Bunge, R. P. & Bunge, M. B. *J. Cell Biol.* **78**, 943–950 (1978).
- Sidman, R. L. & Wessells, N. K. *Expl. Neurol.* **48**, 237–251 (1975).

Functional synapses are established between ciliary ganglion neurones in dissociated cell culture

Joseph F. Margiotta & Darwin K. Berg

Department of Biology, University of California, San Diego, La Jolla, California 92093, USA

The formation of specific synaptic connections is a central part of development in the nervous system. The parasympathetic ciliary ganglion of the chick is a useful population of neurones for studying synaptogenesis both because its development has been well characterized *in vivo*¹ and because its neurones can be maintained and examined in long-term dissociated cell culture^{2,3}. The ciliary ganglion contains two classes of neurones: choroid neurones that innervate smooth muscle in the choroid layer of the eye, and ciliary neurones that innervate striated muscle in the iris and ciliary body. Both classes of neurone are cholinergic and both receive excitatory cholinergic input from preganglionic neurones in the accessory oculomotor nucleus. Ciliary ganglion neurones do not seem to innervate each other *in vivo*^{4,5}, even though they have matching neurotransmitter and receptor types. In cell culture, the neurones acquire high levels of choline acetyltransferase activity, form cholinergic synapses on skeletal myotubes when present in the cultures, and have significant levels of acetylcholine (ACh) sensitivity^{2,6}, as they do *in vivo*. We report here that cholinergic synaptic transmission does occur between the neurones in cell culture. These results indicate that ciliary ganglion neurones can innervate each other, and suggest that additional constraints exist *in vivo* to prevent them from doing so.

Dissociated ciliary ganglion neurones were obtained from 7½-day-old chick embryos and, unless otherwise indicated, were grown in 35-mm dishes with skeletal myotubes as previously described². In some cases the neurones were grown alone, using a collagen substratum coated with fibroblast material². Culture medium was as previously described for nerve-muscle cultures² except that some cultures of neurones alone received medium with 3% (v/v) embryonic chick eye extract instead of 5% (v/v) whole embryo extract². After 5–14 days the neurones were

examined using intracellular recording techniques^{2,6}; micro-electrodes were filled with 1 M potassium acetate and had tip resistances of 80–200 MΩ. Recordings were accepted only if impaled neurones had resting potentials exceeding –45 mV and fired impulses in response to intracellular stimulation.

Intracellular recordings from the neurones revealed discrete spontaneous depolarizations that ranged in amplitude over 0.5–15 mV and in half rise time over 0.6–1.2 ms, and appeared to decay exponentially (Fig. 1). The frequencies of such depolarizations among neurones varied from 2 to 300 per min and persisted throughout recordings that lasted up to 45 min. At high frequencies the depolarizations often summed and occasionally triggered action potentials in the neurones (Fig. 1). Over 90% of the neurones grown with myotubes displayed such depolarizations whereas only 42% of the neurones grown alone did so (Table 1).

Pharmacological experiments indicated that the spontaneous depolarizations were cholinergic in origin. Perfusion of the cultures with 25 μM (+)tubocurarine (TC) completely abolished the depolarizations in all neurones within 5 min; partial recovery of the spontaneous activity was obtained 4 min after washout of the drug was initiated (Fig. 2). After 30 min of washout, the occurrence of depolarizing potentials among randomly selected, treated neurones was indistinguishable from that obtained in control cultures (Table 1).

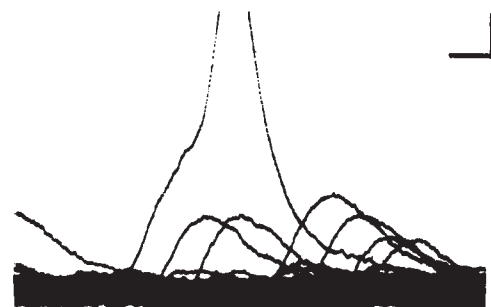


Fig. 1 Spontaneous depolarizations recorded in a ciliary ganglion neurone grown with myotubes. A number of oscilloscope traces have been superimposed showing about 15 depolarizations. In one case a depolarization has triggered an action potential. Resting potential: –61 mV. Calibration bars: vertical, 5 mV; horizontal, 2 ms.

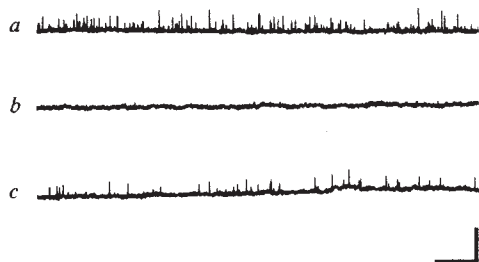


Fig. 2 Reversible blockade of the spontaneous depolarizations by (+)tubocurarine. *a*, Spontaneous depolarizations in a ciliary ganglion neurone grown with myotubes; *b*, blockade of the depolarizations in the same neurone 4 min after the initiation of perfusion with 25 μ M TC; *c*, recovery of the depolarizations in the same neurone 4 min after washout of the TC was initiated by perfusion with normal medium. Resting potential: -60 mV. Calibration bars: vertical, 20 mV; horizontal, 2 s.

The spontaneous depolarizations were also blocked by the α -neurotoxin Bgt 3.1 which has previously been shown to block the ACh sensitivity of ciliary ganglion neurones in cell culture^{6,7}. Neurones incubated with 20 nM Bgt 3.1 for 1 h at 37 °C were never seen to display the spontaneous depolarizations (Table 1). Recovery of the responses was observed after removal of the Bgt 3.1 (Table 1). During TC or Bgt 3.1 treatment, the neurones remained able to fire action potentials when stimulated intracellularly, and had resting potentials not significantly different from those recorded in normal medium ($P > 0.05$, Student's *t*-test).

α -Bungarotoxin (Bgt 2.2), which inhibits ACh receptors on muscle cells but not ACh receptors on ciliary ganglion neurones⁶, did not block the spontaneous depolarizations observed in the neurones. After a 1-h incubation in 10 nM Bgt 2.2, most of the neurones tested still displayed the spontaneous potentials (Table 1). The Bgt 2.2 treatment was effective, however, at blocking muscle ACh receptors, because the normally twitching myotubes in the nerve-muscle cultures were quiescent after the incubation with Bgt 2.2. Intracellular recording demonstrated that the synaptic potentials normally observed in the myotubes were absent. These results indicate

Table 1 Incidence and pharmacological blockade of spontaneous depolarizations in ciliary ganglion neurones

Culture type	Drug treatment	% Neurones with spontaneous depolarizations
Nerve-muscle	None	92 (60, 12)
Nerve	None	42 (24, 8)
Nerve-muscle	TC	0 (13, 6)
	TC removed	93 (14, 6)
Nerve-muscle	Bgt 3.1	0 (13, 3)
	Bgt 3.1 removed	83 (6, 2)
Nerve-muscle	Bgt 2.2	82 (11, 2)

Ciliary ganglion neurones were grown with skeletal myotubes (nerve-muscle) or grown alone (nerve), and incubated with the indicated drug for 1 h at 37 °C, and then examined in the same conditions for spontaneous depolarizations as described in the text for a period of at least 3 min per neurone. The mean resting potential was -56 ± 1 mV ($\pm$ s.e., $n = 80$). The proportion of neurones with spontaneous depolarizations is presented as a per cent of the neurones tested; the number of neurones tested and the number of cultures sampled are shown in parentheses. Drug concentrations were TC, 25 μ M; Bgt 3.1, 20 nM; and Bgt 2.2, 10 nM. For one of the TC cultures, 25 μ M hexamethonium was included and the TC concentration was 100 μ M. Recovery from TC treatment was followed in the same cultures used to test the effects of TC: at the end of the TC test period the cultures were rinsed and incubated for 0.5 h at 37 °C and then re-examined (TC removed). As for recovery from Bgt 3.1, cultures previously tested in Bgt 3.1 were rinsed and incubated for 2 h at 37 °C and then re-examined (Bgt 3.1 removed). In all cases the cultures were also tested before drug treatment to verify that the neurones displayed the expected incidence ($\sim 90\%$) of spontaneous depolarizations.

that the spontaneous depolarizations observed in the neurones are not in some way dependent on neurogenic muscle contraction, a conclusion also implied by the presence of the depolarizations in neurones grown alone.

The pharmacological experiments indicate that the spontaneous depolarizations recorded in the neurones represent cholinergic excitatory postsynaptic potentials. They can be reversibly blocked by TC and Bgt 3.1, and are not blocked by Bgt 2.2. A similar pharmacology has been described for ACh receptors on the neurones⁶. The implication is that ciliary ganglion neurones establish synapses on each other in cell culture and that neurotransmitter release can occur spontaneously or be evoked by spontaneous action potentials in the presynaptic neurone. Synapses between ganglionic neurones have also been described for rat sympathetic neurones in cell culture when the cells are induced to become cholinergic^{8,9}.

Morphological evidence for chemical synapses between ciliary ganglion neurones in culture has been reported previously, but stimulation of the neurones failed to evoke synaptic transmission¹⁰. We routinely test neurones for electrical excitability by using intracellular stimulation to trigger action potentials. We have occasionally seen depolarizations produced in a neurone following action potentials elicited in the same neurone. The evoked depolarizations were phase-locked to the action potentials with a fixed time delay and, in the case of one neurone, were adequate to elicit a second action potential (Fig. 3). Although we have not demonstrated that the evoked depolarizations can be blocked by cholinergic antagonists, they do have the range of amplitude and half rise time described above for the synaptic potentials. Studies with tetrodotoxin (TTX) support the possibility that evoked transmission does occur between the neurones. Perfusion of nerve-muscle cultures with 10 μ M TTX for 2 min completely abolished the spontaneous synaptic potentials (eight neurones, three cultures); rinsing the cultures with medium lacking TTX for 4 min resulted in the re-appearance of synaptic potentials in four out of five neurones tested (three cultures). These results are consistent with the synaptic potentials being elicited by spontaneous TTX-sensitive action potentials which we frequently observed while recording from the neurones (unpublished observations). To determine whether evoked transmission can be reliably obtained at synapses capable of spontaneous transmission, however, it will be necessary to identify transmitting neurone pairs and to test them with intracellular stimulation.

The difference in the proportion of neurones showing synaptic potentials in nerve compared with nerve-muscle cultures (Table 1) could reflect differences in the frequency of action

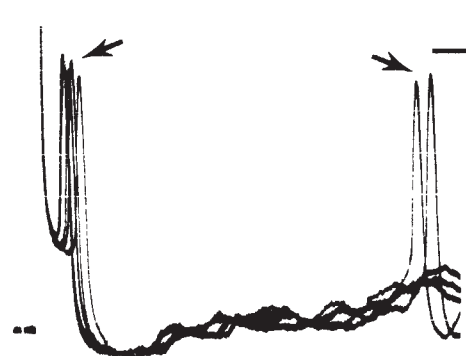


Fig. 3 Recordings from a neurone grown with myotubes showing evoked depolarizations occurring with a fixed time delay after an action potential. Five superimposed oscilloscope traces are shown. In each case a brief (0.1 ms) intracellular current pulse was used to trigger an action potential (arrow, left) in the impaled neurone. After a delay of ~ 10 ms, a series of depolarizations was observed. In two of the five trials the depolarizations triggered a second action potential (arrow, right). The nearly vertical line at the far left marks the stimulus artefact. Resting potential: -58 mV. Calibration bars: vertical, 10 mV; horizontal, 5 ms.

potentials in presynaptic neurones or differences in the number or efficacy of synapses in the two kinds of cultures. The presence of myotubes as a synaptic target for the neurones clearly does not prevent the neurones from also forming synapses on each other.

Although it remains possible that ciliary ganglion neurones transiently innervate each other at early stages *in vivo*, it seems clear that they normally do not do so in the adult chick or pigeon^{4,5}. Recently it has been shown that parasympathetic neurones of the frog cardiac ganglion which also do not normally innervate each other, can do so *in vivo* when deprived of their normal preganglionic input; reinnervation of the neurones by the preganglionic fibres leads to suppression of the synapses between ganglionic neurones¹¹. The formation of synapses between ciliary ganglion neurones in culture may also reflect the absence of their normal preganglionic fibres.

This work was supported by grants from the NIH (NS 12601 and MH 16109) and from the Muscular Dystrophy Association. We thank Dr Martin Smith for providing some of the data obtained with Bgt 3.1, Marie Ryder for providing technical assistance, and Dr Nicholas Spitzer for comments on the manuscript.

Received 24 August 1981; accepted 11 January 1982.

- Landmesser, L. & Pilar, G. *J. Physiol., Lond.* **222**, 691–713 (1972); *J. Physiol., Lond.* **241**, 737–749 (1974); *J. Cell Biol.* **68**, 357–374 (1976); *Fedn Proc.* **37**, 2016–2022 (1978).
- Nishi, R. & Berg, D. K. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5171–5175 (1977); *Nature* **277**, 232–234 (1979); *J. Neurosci.* **1**, 505–513 (1981).
- Tuttle, J. B., Suszkiw, J. B. & Ard, M. *Brain Res.* **183**, 161–180 (1980).
- Marwitt, R., Pilar, G. & Weakly, J. N. *Brain Res.* **25**, 317–334 (1971).
- Giacobini, E., Pilar, G., Suszkiw, J. & Uchimura, H. *J. Physiol., Lond.* **286**, 233–253 (1979).
- Ravdin, P. M. & Berg, D. K. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2072–2076 (1979).
- Ravdin, P. M., Nitkin, R. M. & Berg, D. K. *J. Neurosci.* **1**, 849–861 (1981).
- O'Laigue, P. H., Obata, K., Claude, P., Furshpan, E. J. & Potter, D. D. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3602–3606 (1974).
- Furshpan, E. J., MacLeish, P. R., O'Laigue, P. H. & Potter, D. D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4225–4229 (1976).
- Pilar, G. & Tuttle, J. *Soc. Neurosci. Abstr.* **4**, 593 (1978).
- Sargent, P. B. & Dennis, M. J. *Dev. Biol.* **81**, 65–73 (1981).

Growth factors adherent to cell substrate are mitogenically active *in situ*

J. C. Smith*, J. P. Singh, J. S. Lillquist, D. S. Goon & C. D. Stiles

Laboratory of Tumor Biology, Sidney Farber Cancer Institute, and Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

We report here that platelet-derived growth factor (PDGF)^{1–4} and pituitary-derived fibroblast growth factor (FGF)⁵ bind rapidly and tenaciously to tissue culture plates either coated with collagen (a major component of the extracellular matrix) or uncoated. Surface-bound PDGF resists extraction with non-ionic detergents, organic solvents and chaotropic solvents; moreover it is biologically active *in situ*. Our finding that immobilized PDGF is nonetheless mitogenic lends support to other observations which indicate that PDGF functions outside the cell via production of a stable intracellular second signal⁶. In a broader context, the observations suggest that one role of the extracellular matrix *in vivo* is to sequester biologically active molecules like PDGF to provide a localized and persistent stimulation.

PDGF was purified to homogeneity and radioiodinated as described in Table 1 legend. Binding studies with ¹²⁵I-PDGF (see Table 1 legend) showed that the radiolabelled material binds to specific receptors on the surface of confluent BALB/c-3T3 cell monolayers. We and others⁷ have found that most of

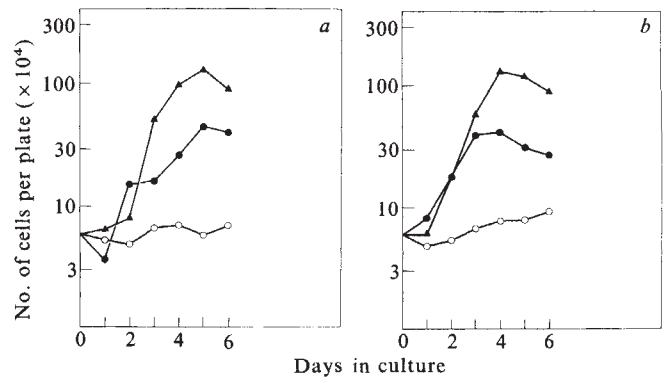


Fig. 1 PDGF-treated plates support continued proliferation of BALB/c-3T3 cells in medium containing platelet-poor plasma. Tissue culture plates (35 mm), either coated with collagen or uncoated, were exposed to 500 units ml⁻¹ of heat-treated PDGF³ at 37 °C for 4 h. Control plates were not treated with PDGF. BALB/c-3T3 cells (6 × 10⁴) were seeded on to each plate in medium containing 5% platelet-poor plasma and cell numbers were determined daily; cells were fed every other day. For comparison, the growth of cells on untreated plates in medium containing 5% calf serum is shown. *a*, Collagen-coated plates: ●, 5% plasma, PDGF-treated plates; ○, 5% plasma, untreated plates; ▲, 5% serum, untreated plates. *b*, Uncoated plates: symbols as in *a*.

the ¹²⁵I-PDGF bound to confluent monolayer cell cultures is cell associated. In the absence of a cell monolayer, however, ¹²⁵I-PDGF binds to the surface of empty tissue culture plates in a non-saturable manner. In the experiments summarized in Table 1, ¹²⁵I-PDGF (20 ng ml⁻¹) was placed over the surface of empty microtitre tissue culture wells and incubated at 37 °C for 1 h. In these conditions about 1 ng of PDGF becomes associated with the surface of the culture dish, and this surface-bound PDGF resists extraction with various strong solvents. We found that boiling SDS is partially effective in extracting the bound PDGF, and pronase solubilizes ~70% of the radioactivity. From the molecular weight of PDGF (33,000–35,000)^{3,4} it can be calculated that the empty tissue culture wells bind between 6 × 10¹⁰ and 9 × 10¹⁰ molecules of PDGF per cm². Although the absolute amount of PDGF bound is small, a 3T3 cell (spread surface area 2 × 10⁻⁵ cm²) plated on to a culture dish pretreated with PDGF could be presented with between 3 × 10⁵ and 4.5 × 10⁵ molecules of PDGF from below. This is well in excess of the number of molecules needed to saturate cellular PDGF receptors (see Table 1 legend).

When quiescent 3T3 cells were plated out in culture dishes that had been pretreated with PDGF, DNA replication was induced in the absence of any PDGF added to the medium (Table 2), thus PDGF bound to the surface of tissue culture dishes is mitogenically active. F6F, which is a functional analogue of PDGF for stimulating the growth of BALB/c-3T3 cells⁸, also retained its mitogenic activity when bound to tissue culture dishes (Table 2). The concentrations of pure PDGF or partially pure FGF which delivered a mitogenic quantity of growth factor to the surface of empty tissue culture plates (Table 2) were of the same order of magnitude as those used to stimulate cell growth in conventional conditions^{3,5}. The mitogenic activity of PDGF fixed to the surface of tissue culture plates resisted extraction with 6 M urea, methanol, ethanol and glycine-EDTA (pH 10) as well as Triton X-100 (Table 3), but was totally abolished by reduction with 2-mercaptoethanol or digestion with proteases, as is the activity of soluble PDGF³. In agreement with the data of Table 1, boiling 1% SDS partially removed or inactivated the surface-bound PDGF.

We demonstrated that surface-bound PDGF is biologically active *in situ* in the following way. One-half of an empty tissue culture plate was exposed to PDGF by tilting the plate during treatment. The plate was washed with phosphate-buffered saline (PBS) and then 3T3 cells were seeded over the entire plate at a density of 7 × 10⁴ cm⁻² in medium containing 5% plasma

*Present address: Imperial Cancer Research Fund, Mill Hill Laboratories, Burtonhole Lane, London NW7 1AD, UK.

and ^3H -TdR. The culture plate was incubated for 42 h, fixed and processed for autoradiography. The amount of labelling on the treated side of the plate was 42%, that on the untreated side 7%. Hence, the mitogenic activity of PDGF-treated plates is predominantly a local effect.

Surface-bound PDGF supports repeated cycles of cell growth and division as well as localized induction of DNA synthesis. We compared proliferation of BALB/c-3T3 cells on four substrates (plastic, collagen, plastic pre-treated with PDGF, and collagen pre-treated with PDGF) in medium containing 5% plasma (Fig. 1). Cells maintained on plastic or collagen proliferated only

Table 1 Binding of ^{125}I -PDGF to the surface of tissue culture dishes

PDGF-treated plates extracted with:	c.p.m. ^{125}I remaining on surface of dish
PBS	10,806
Triton X-100 (0.5%)	11,389
Methanol (95%)	12,896
Ethanol (95%)	12,362
Urea (6 M)	15,495
SDS (1%) (100 °C)	7,963
Pronase (2 mg ml $^{-1}$)	3,760

PDGF was purified to homogeneity by a modification of our previous method³, then radioiodinated to a specific activity of 10,000 d.p.m. ng $^{-1}$ by a modification of the Hunter and Greenwood procedure¹⁸. The ^{125}I -PDGF retained its mitogenic activity. Studies of binding specificity to BALB/c-3T3 cells (performed at 4 °C) indicated that confluent monolayer cultures of BALB/c-3T3 possess $\sim 8 \times 10^4$ PDGF receptors per cell. Nanogramme per ml quantities of unlabelled PDGF completely displaced the ^{125}I -PDGF from these surface receptors whereas microgramme per ml quantities of epidermal growth factor or insulin had no effect. To measure the binding of ^{125}I -PDGF to the cell substratum, empty microtitre tissue culture wells (0.3 cm 2 surface area) were incubated with 40,000 c.p.m. of ^{125}I -PDGF in 200 μl Dulbecco's modified Eagle's medium (DMEM) with 1 mg ml $^{-1}$ of unlabelled bovine serum albumin as carrier. After 1 h at 37 °C the ^{125}I -PDGF solution was aspirated. The microtitre plates were then extracted at 37 °C for 1 h with the indicated solvents (except for SDS which was a 1 min extraction at 100 °C). Following extraction the solvents were aspirated. The microtitre plates were again washed extensively with PBS and dried. The bottom of each microtitre plate was punched out and placed in a γ counter for determination of bound ^{125}I -PDGF.

Table 2 Mitogenic activity of surface-bound PDGF and FGF

Empty plates pretreated with PDGF		Empty plates pretreated with FGF	
[PDGF] (ng ml $^{-1}$)	% Labelled nuclei	[FGF] (ng ml $^{-1}$)	% Labelled nuclei
40	94.9	1,000	91
20	95.2	300	65
10	83.2	100	58
2.5	65.3	30	8
0	24.5	0	2

The mitogenic activity of surface-bound PDGF and FGF was detected by induction of DNA synthesis in quiescent 3T3 cells. Empty microtitre tissue culture wells were precoated with collagen (collagen dispersion TD-150; Ethicon, Massachusetts) to facilitate rapid cell attachment. The collagen-coated plates were then pretreated with the indicated concentrations of electrophoretically pure PDGF (prepared by a modification of a previous protocol³) or partially purified pituitary FGF (KOR Biochemicals, Massachusetts) at 37 °C for 4 h. The plates were rinsed twice with PBS and further extracted at 37 °C for 1 h with 0.5% Triton X-100 as described in Table 1 legend. Quiescent density-arrested BALB/c-3T3 cells were collected by trypsin digestion and immediately replated at a density slightly greater than their confluent density ($\sim 7 \times 10^4$ cells cm $^{-2}$) into the treated microtitre culture wells. The cultures were incubated at 37 °C for 36 h in medium containing ^3H -thymidine (5 μCi ml $^{-1}$) and 5% platelet-poor plasma (which lacks PDGF 1,2). The cultures were then fixed and processed for autoradiography. Per cent labelled cells was determined for each well.

Table 3 Mitogenic activity of surface-bound PDGF after extraction with various solvents and enzymes

PDGF-treated plates extracted with:	% Labelled nuclei $\pm$ s.d.
Triton X-100 (0.5%)	66 $\pm$ 20 (n = 8)
Glycine (10 mM; pH 10.0)	88 $\pm$ 5 (3)
Methanol	92 $\pm$ 5 (2)
Ethanol	84 $\pm$ 4 (2)
Urea (6 M)	68 $\pm$ 17 (3)
RNase (30 μg ml $^{-1}$)	77 $\pm$ 5 (2)
DNase (10 μg ml $^{-1}$)	74 $\pm$ 5 (2)
Acetic acid (1 M) (pH 2.3)	46 $\pm$ 12 (3)
SDS (1%; 100 °C)	40 $\pm$ 32 (3)
Urea (6 M) + 2-mercaptoethanol (10%)	10 $\pm$ 8 (2)
Pronase (2 mg ml $^{-1}$)	0 $\pm$ 5 (3)
Trypsin (1,000 U ml $^{-1}$)	19 $\pm$ 17 (2)

The surface of empty tissue culture plates was exposed to partially purified PDGF³ in DMEM at 37 °C for 3–4 h. The PDGF concentrations used were in the range 50–500 units ml $^{-1}$ where 1 unit of activity = ~ 0.2 ng PDGF³. Control plates were treated with DMEM alone. After 3–4 h the solutions were aspirated. The empty wells were then washed and extracted with solvents as described in Table 1 legend. Surface-bound, biologically active PDGF was detected by induction of DNA synthesis in quiescent 3T3 cells as described in Table 2 legend. The nuclear labelling index of cells cultured on the treated plates is shown after subtracting the background of labelled nuclei on untreated plates; the background labelling ranged from 0 to 25% in individual experiments.

poorly in 5% plasma but grew rapidly for 3–5 days on substrate pre-treated with PDGF. Cells cultured with 5% plasma on PDGF-treated plates grew to slightly lower final densities than cells cultured conventionally in serum-containing medium; however, the cells cultured conventionally received three separate PDGF treatments (through the change of serum-containing medium on alternate days) whereas the cells cultured on PDGF-treated plates were subject to only one PDGF treatment. We have obtained similar results by treating collagen-coated plates with pituitary FGF.

We conclude that PDGF and FGF adhere tenaciously to the surface of tissue culture plates; from this position, they can stimulate proliferation of cells maintained in platelet-poor plasma. These results may have a bearing on tissue culture analyses of extracellular matrix materials, and their role in controlling cell growth and function^{9–11}. The findings may also be relevant to the well characterized 'anchorage requirement' for fibroblast growth *in vitro*^{12–14}. It remains to be determined whether this has physiological relevance *in vivo*.

The induction of cell division by surface-bound PDGF is a highly efficient process as quantified by dose-response analysis, amount of PDGF consumed and duration of effect. If efficiency can be taken as a sign of biological relevance, these data suggest that fibroblast-like cells are intended to respond to PDGF presented in this manner. A functional interaction between PDGF and the cellular substratum *in vivo* can easily be rationalized. PDGF is a connective tissue mitogen which is also chemotactic for smooth muscle cells¹⁵; it is believed to have a role in wound healing and maintenance of the vascular system and may be involved in the genesis of atherosclerosis^{2,16,17}. If PDGF adheres to collagen surrounding the smooth muscle cells at the site of a wound, the mitogenic effect might be considerably prolonged. A gradient of PDGF concentration so produced (high near the wound, lower further away) might provide a directional cue for migrating cells. More generally, the adhesion of biologically active polypeptides to the extracellular matrix might serve to guide cell migration during embryogenesis or even provide a 'memory' of past developmental signals.

This research was supported by the NCI. J.C.S. is a NATO postdoctoral fellow. C.D.S. is supported by a Faculty Research Award from the American Cancer Society. We thank Lynne Dillon for typing the manuscript.

Received 6 July 1981; accepted 14 January 1982.

- Kohler, N. & Lipton, A. *Expl Cell Res.* **87**, 297–301 (1974).
- Ross, R., Glomset, J., Kariya, B. & Harker, L. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1207–1211 (1974).
- Antoniades, H. N., Scher, C. D. & Stiles, C. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1809–1813 (1979).
- Heldin, C. H., Westermark, B. & Wasteson, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3722–3726 (1979).
- Gospodarowicz, D. *J. biol. Chem.* **250**, 2515–2520 (1975).
- Smith, J. C. & Stiles, C. D. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4363–4367 (1981).
- Heldin, C. H., Westermark, B. & Wasteson, A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3664–3668 (1981).
- Stiles, C. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1279–1283 (1979).
- Gospodarowicz, D., Vlodavsky, I., Greenburg, G. & Johnson, L. K. *Cold Spring Harb. Conf. on Cell Proliferation*, 6 (eds Sato, G. H. & Ross, R.) 561–592 (Cold Spring Harbor Laboratory, New York, 1979).
- Gospodarowicz, D. & Ill, C. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2726–2730 (1980).
- Rojkind, M. *et al. J. Cell Biol.* **87**, 255–263 (1980).
- Macpherson, I. & Montaignier, L. *Virology* **23**, 291–294 (1964).
- Stoker, M. *Nature* **218**, 234–238 (1968).
- Folkman, J. & Moscona, A. *Nature* **273**, 345–349 (1978).
- Gotendorst, G. R., Seppa, H. E. J., Kleinman, H. K. & Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Ross, R. & Vogel, A. *Cell* **14**, 203–210 (1978).
- Scher, C. D., Shepard, R. C., Antoniades, H. N. & Stiles, C. S. *Biochim. biophys. Acta* **560**, 217–241 (1979).
- Hunter, W. M. & Greenwood, F. C. *Nature* **194**, 495–496 (1962).

Selective effect of feline leukaemia virus on early erythroid precursors

D. Onions*, O. Jarrett*, N. Testa†, F. Frassoni† & S. Toth*

* Department of Veterinary Pathology, University of Glasgow, Bearsden Road, Glasgow G61 1QH, UK

† Paterson Laboratories, Christie Hospital and Holt Radium Institute, Manchester M20 9BX, UK

Feline leukaemia virus (FeLV), a horizontally transmitted retrovirus, is the aetiological agent of feline lymphosarcoma and leukaemia^{1–3}. In addition, the establishment of persistent FeLV viraemia in cats has been associated with the induction of a wide range of non-neoplastic conditions including immunosuppression, reproductive disorders and anaemias^{4,5}. FeLV isolates from cats may be divided into three subgroups, A, B and C, based on their interference patterns^{6,7}. In natural isolates, FeLV-C is always found in association with subgroups A or AB but the subgroup C component may be separated as a non-defective virus capable of establishing a persistent viraemia in neonatal cats. Pure red cell hypoplasia (PRCH) is one of three primary forms of FeLV-associated anaemia recognized in cats^{8,9} and we report here that three new biologically cloned FeLV-C isolates will reproduce this disease. We have demonstrated that infection of cats with FeLV-C/Sarma results in rapid depletion of early erythroid precursors (burst-forming units—erythroid; BFU-E) while the proportion of the myeloid precursor (granulocyte-macrophage colony-forming cells; GM-CFC) in the bone marrow remains normal.

In a survey of FeLV subgroups in naturally infected, viraemic cats, FeLV-C was not observed in 98 clinically normal animals but was isolated from 9 of the 32 (28%) anaemic cats examined. Subgroup C viruses from three of these anaemic cats were biologically cloned and used to infect neonatal kittens as described in Table 1 legend. Each isolate produced a PRCH that was clinically evident 3–6 weeks after infection. The pathology of the disease produced by these isolates resembled that previously described for FeLV-C/Sarma¹⁰.

To determine the effect of subgroup C viruses on haematopoietic precursor cells, 3.5×10^5 focus-inducing units (FIU) of FeLV-C/Sarma were inoculated intraperitoneally, within 24 h of birth, into specific pathogen-free randomly bred cats. A control group of uninfected cats was established from the same colony and members of both groups were bled weekly to determine their blood constituents and virus status¹¹. We detected viraemia in all the FeLV-infected cats except cat 7.

The control cats showed normal feline blood development¹² in that the haematocrit rose from a low value at 4 weeks until the adult level was attained at ~12 weeks (Fig. 1). In the infected cats, however, the rise in the haematocrit was not sustained but began to fall from 7 weeks onwards. Individual FeLV-C-infected cats displayed episodes of splenic erythroid regeneration with the release of normoblasts into the circulation. These events were associated with a temporary slight rise in the individual's haematocrit before it fell to a new low level. Cat 7, which was not detectably viraemic, displayed normal blood development throughout the experiment.

Starting 3 weeks after infection and at intervals thereafter, the femoral bone marrow of the cats was assayed for the early erythroid precursor (BFU-E)¹³ and the neutrophil macrophage precursor (GM-CFC; refs 14, 15) (Table 2). The numbers of BFU-E and GM-CFC in all the control cats were within the range previously established for cats 3–12 months old (data not shown). Examination of the FeLV-C-infected cats, however, indicated that a dramatic reduction in the numbers of BFU-E cells had occurred as early as 3 weeks after infection while the incidence of GM-CFC was normal (Table 2). A reduction in BFU-E numbers in FeLV-C-infected cats has also been observed by others (J. Adamson, personal communication).

Both control and FeLV-C-infected cats exhibited splenic extramedullary haematopoiesis (EMH) which is a normal feature of feline haematopoiesis in cats of this age range. This process may have accounted for the normal haematological status of cat 7, which was not detectably viraemic but which showed extensive reduction of BFU-E in the bone marrow. The pattern of reduced BFU-E and normal GM-CFC levels was seen in the infected cats throughout the experiment, except

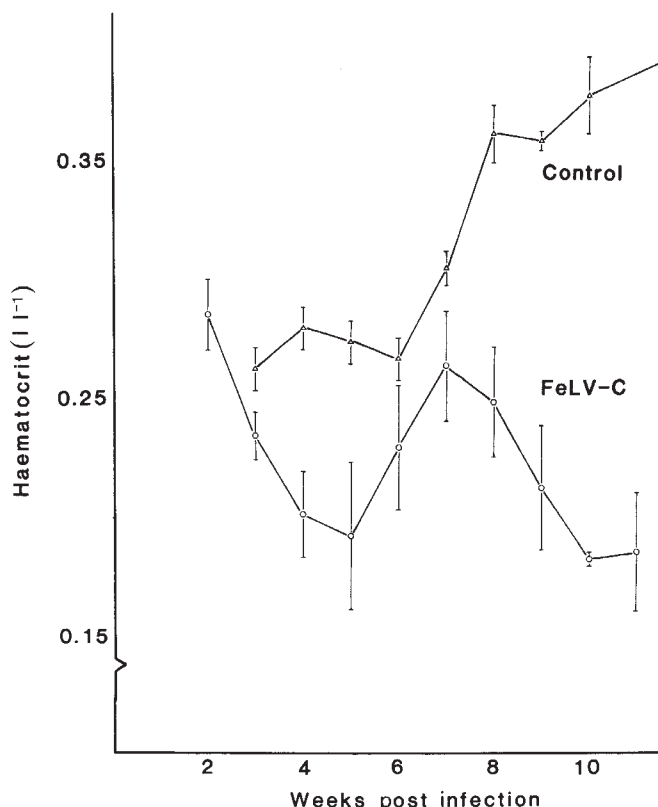


Fig. 1 Mean haematocrits ($\pm$ s.e.) of control uninfected and FeLV-C/Sarma-infected kittens. A control group of 6 newborn kittens and 10 FeLV-C/Sarma-infected newborn kittens were established from the same colony. A separate group of three control and eight of the FeLV-C/Sarma-infected kittens were used in the colony assays described in Table 2 (cats 1–11). The kittens were bled weekly by jugular venepuncture and a 1-ml blood sample was used to establish the haematological values and the presence or absence of plasma viraemia as previously described¹¹. A comparison of the haematocrits of the FeLV-C-infected and control group by Wilcoxon's two-sample rank test indicated significant differences (0.01 level of T) at weeks 5, 8 and 9. From week 10 onwards only two kittens remained in the FeLV-C-infected group, precluding a statistical comparison.

Table 1 Development of the haematocrit and total leukocyte count in FeLV-C-infected cats

Age (weeks)	Haematocrit ($11^{-1} \pm \text{s.e.}$)				Leukocyte count ($10^9 \text{ l}^{-1} \pm \text{s.e.}$)	
	Control	FeLV-C/ FS246	FeLV-C/ FA27	FeLV-C/ FZ215	FeLV-C/ Sarma	Control
3	0.262 $\pm$ 0.009	0.250 $\pm$ 0.007	0.254 $\pm$ 0.008	0.197 $\pm$ 0.016	14.0 $\pm$ 1.6	14.8 $\pm$ 3.1
4	0.279 $\pm$ 0.009	0.234 $\pm$ 0.003	0.196 $\pm$ 0.005	0.209 $\pm$ 0.009	14.9 $\pm$ 2.1	14.7 $\pm$ 2.6
5	0.274 $\pm$ 0.009	ND	0.205 $\pm$ 0.013	0.149 $\pm$ 0.017	17.6 $\pm$ 5.3	15.7 $\pm$ 1.9
6	0.276 $\pm$ 0.005	0.184 $\pm$ 0.006	0.213 $\pm$ 0.015	0.127 $\pm$ 0.012	15.3 $\pm$ 1.7	13.2 $\pm$ 2.5
7	0.304 $\pm$ 0.007	0.202 $\pm$ 0.010	0.256 $\pm$ 0.090	0.129 $\pm$ 0.015	12.7 $\pm$ 2.7	12.4 $\pm$ 1.8
8	0.394 $\pm$ 0.026	0.205 $\pm$ 0.015	0.258 $\pm$ 0.011	0.125 $\pm$ 0.023	21.0 $\pm$ 4.5	18.3 $\pm$ 2.9
9	0.358 $\pm$ 0.004	0.213 $\pm$ 0.026	ND	0.123 $\pm$ 0.026	16.8 $\pm$ 3.2	14.5 $\pm$ 1.9
10	0.376 $\pm$ 0.017	0.188 $\pm$ 0.028	ND	0.092 $\pm$ 0.024	17.4 $\pm$ 1.5	19.1 $\pm$ 2.2
11	ND	0.185 $\pm$ 0.022	0.258 $\pm$ 0.014	—	21.2 $\pm$ 1.9	24.7 $\pm$ 3.0
12	0.394 $\pm$ 0.009	0.187 $\pm$ 0.028	0.248 $\pm$ 0.024	—	—	—
13	ND	0.185 $\pm$ 0.023	0.231 $\pm$ 0.030	—	—	—
14	0.358 $\pm$ 0.011	0.172 $\pm$ 0.017	0.182 $\pm$ 0.040	—	—	—
15	ND	0.159 $\pm$ 0.019	0.180* $\pm$ 0.044	—	—	—
16	0.371 $\pm$ 0.008	0.136 $\pm$ 0.017	0.197 $\pm$ 0.050	—	—	—

Three new FeLV-C isolates, FZ215, FS246 and FA27, were biologically cloned, checked for subgroup-C specificity by interference assay and the titre determined as described previously¹¹. The pathogenicity of these isolates was determined by intraperitoneal inoculation to groups of four 1-day-old kittens with either 3.4×10^5 FIU of FZ215, 1.0×10^4 FIU of FS246 or an undetermined titre of FA27 (ref. 23). A group of six age-matched controls remained uninfected. The haematocrit and plasma viraemia were determined each week for 3 weeks. In a separate experiment (described in Fig. 1 legend) day-old kittens were either inoculated with FeLV-C/Sarma or uninfected (controls) and the weekly total leukocyte counts measured. The haematocrit and bone marrow colony assays for this experiment are given in Fig. 1 and Table 2, respectively. ND, not determined. —, Experiment completed.

* Cat removed with a haematocrit of 0.092.

for cat 10 which, when killed at 11 weeks, had reduced GM-CFC. There may have been a limited effect of FeLV-C on GM-CFC but a simple interpretation of this effect was complicated by the occurrence of myelofibrosis and osteosclerosis. Although there was a marked reduction in total bone marrow cellularity in FeLV-C/Sarma-infected cats, both these cats and the controls exhibited a progressive rise in leukocyte numbers (Table 1), typical of kittens of this age¹². There is therefore a clear selective effect of FeLV-C in ablating the precursors of erythroid cells compared with those of myeloid cells and this effect is reflected in the circulating numbers of end cells of each of these compartments.

Myelofibrosis and medullary osteosclerosis have been described as accompanying features of experimentally produced PRCH¹⁶. These features were also seen in the present experiments which used cloned FeLV-C isolates but the extent and time of development of the conditions displayed considerable individual variation. In the experiment with FeLV-C/Sarma, early histopathological signs of myelofibrosis and osteosclerosis were detectable in the infected cats after 5 weeks but were absent from those examined at 7 weeks. At 11 weeks, cat 10 showed a marked development of both medullary osteosclerosis and myelofibrosis not evident in cat 11. It is apparent, therefore, that the anaemia associated with FeLV-C infection is not due to myelophthisis, as the former can occur in the absence of the myelofibrosis-osteosclerosis complex. It is also unlikely that the osteosclerosis and myelofibrosis are secondary consequences of the anaemia since osteosclerosis is not a feature of feline anaemias induced by other agents.

It is possible that some PRCH cases are associated with FeLV-A or FeLV-AB infections. However, the pronounced association of FeLV-C with anaemia and the ability of all four cloned FeLV-C isolates to induce PRCH suggest that this disease is predominantly associated with this subgroup. Furthermore, none of the FeLV-A or FeLV-B isolates tested in pathogenesis experiments have produced PRCH although other forms of anaemia have been reported^{8,11,17}.

The mechanism of FeLV-C-induced depletion of BFU-E is unknown. It does not appear to be associated with a hyporesponsiveness of BFU-E in infected marrow to erythropoietin as levels of erythropoietin five times above the plateau level do not increase numbers of BFU-E observed (Table 2). The

Table 2 Colony forming units of erythroid and myeloid precursors in normal and FeLV-C-infected bone marrow

Age (weeks)	Cat	Status	BFU-E	GM-CFC
3	1	FeLV-C/Sarma	2.0 $\pm$ 1.6	62.0 $\pm$ 9.4
	2	FeLV-C/Sarma	1.0 $\pm$ 1.0	74.0 $\pm$ 13.6
	3	Control	43.0 $\pm$ 8.2	72.0 $\pm$ 8.4
5	4	FeLV-C/Sarma	1.6 $\pm$ 1.6	57.0 $\pm$ 12.2
	5	FeLV-C/Sarma	7.4 $\pm$ 1.8	47.0 $\pm$ 11.4
6	6	Control	30.6 $\pm$ 1.0	61.6 $\pm$ 5.8
7	7	FeLV-C/Sarma	6.0 $\pm$ 1.6	48.0 $\pm$ 11.0
	8	FeLV-C/Sarma	22.0 $\pm$ 4.4	39.4 $\pm$ 5.6
8	9	Control	72.0 $\pm$ 1.6	46.0 $\pm$ 6.2
	10	FeLV-C/Sarma	12.0 $\pm$ 6.2	14.0 $\pm$ 2.4
11	11	FeLV-C/Sarma	2.6 $\pm$ 2.5	31.0 $\pm$ 5.0
	12	Control (0 unit Epo)	0	—
36		Control (0.05 unit Epo)	9.4 $\pm$ 1.8	35.4 $\pm$ 3.7
		Control (0.10 unit Epo)	28.6 $\pm$ 1.4	37.4 $\pm$ 6.8
		Control (0.2 unit Epo)	28.0 $\pm$ 9.8	40.6 $\pm$ 5.8
		Control (0.5 unit Epo)	24.0 $\pm$ 5.4	36.0 $\pm$ 3.0
	13	FeLV-AC (0.1 unit Epo)	1.4 $\pm$ 1.4	50.6 $\pm$ 4.8
		FeLV-AC (0.5 unit Epo)	0.6 $\pm$ 0.6	50.6 $\pm$ 12.4

Colony forming units of erythroid (BFU-E) and myeloid (GM-CFC) precursors per 10^5 bone marrow cells from normal FeLV-C-infected cats. The haematocrits and total leukocyte counts of cats 1–11 are given in Fig. 1 and Table 1, respectively. Cats 12 and 13 were used in a separate experiment to determine the effect of the erythropoietin (Epo) concentration on BFU-E formation; a different batch of anaemic mouse serum was used in these experiments. The technique used to grow erythroid bursts derived from BFU-E of cat bone marrow was modified from that described for other species²⁴. Femoral bone marrow cells suspended in alpha medium (Gibco) plus 2% fetal calf serum (FCS) were cultured in alpha medium containing nucleosides supplemented with 20% FCS (Flow), 10^{-7} M sodium selenite, 3.4×10^{-6} M human transferrin saturated with iron, 10^{-4} M thioglycerol and 1% bovine serum albumin (BSA) in 0.8% 4000 CPS methylcellulose (Dow Chemicals). The source of erythropoietin was anaemic mouse serum (AMS)²⁵ previously titrated against the International Erythropoietin Standard B. The optimum dose for the burst assay in cats 1–11 was 0.4 U Epo per ml of culture. Four aliquots (0.25 ml) each containing 2.5 – 5.0×10^4 cells were plated in wells and incubated at 37 °C in a water-saturated atmosphere of air with 5% CO₂ for 9 days, previously shown to be optimum for the full growth of bursts. A more detailed description of the technique will be reported elsewhere. Colonies of granulocytes and macrophages, derived from their progenitors, GM-CFC, were scored in the same cultures used to determine the numbers of bursts. Previous experiments showed that normal mouse serum or AMS provide an adequate stimulus for the growth of GM-CFC. In these experiments, the amount of AMS used in the erythroid cultures also provided maximum stimulation for the assay of GM-CFC.

decrease in BFU-E may be due to immune elimination of infected cells, to an effect on accessory cells required for erythropoiesis, or a direct block of differentiation of early erythroid precursors. Certain of the adult cases of pure red cell aplasia in man have been postulated to be an autoimmune condition mediated by antibody directed against either erythropoietin or erythroid precursors^{18,19}. Immune elimination of infected precursors in the cat, following viral antigen expression on the surface of infected cells, seems unlikely. This follows from an examination of bone marrow smears for the presence of FeLV group-specific antigen, p30, by the indirect immunoperoxidase technique, which indicated that widespread infection of myeloid cells occurred but the incidence of GM-CFC (Table 2) and the numbers of circulating leukocytes were normal (Table 1). Furthermore, elimination of infected cells is associated with the production of neutralizing antibody and with the consequent abrogation of the infection, while the establishment of clinical anaemia is associated with a persistent infection and the absence of neutralizing antibody. Lymphoid cell infiltrations have been observed in the marrow of humans having pure red cell aplasia¹⁹ and a lymphoid-dependent suppression of haematopoiesis has been suggested for some cases of aplastic anaemia²⁰. In the FeLV-C-infected cats, however, we did not detect any extensive lymphoid infiltration of the marrow. An effect of FeLV-C on accessory cells cannot be excluded. These cells could either represent cells forming the local microenvironment of the bone marrow or T cells which may promote erythropoiesis²¹. In our experiments a gross effect on the T-cell compartment was noted in cat 11 which had an extensive thymic atrophy, while the other FeLV-C infected cats showed no thymic depletion. In other FeLV-C experiments^{8,10} thymic atrophy has been a major finding, but FeLV of other subgroups induce this change without producing anaemia. However, a more subtle modulation of T-cell subsets may have occurred in the FeLV-C-infected cats which affected their capacity to promote or inhibit erythropoiesis.

Any hypothesis explaining the findings presented here must account for the remarkable selective effect of FeLV-C in producing a red cell hypoplasia but with normal levels of GM-CFC and circulating leukocytes. An effect of FeLV-C on the accessory cells required for erythropoiesis could be consistent with these observations but may also be explained by a mechanism that involves a direct effect of the virus on early erythroid precursor cells. The latter hypothesis is consistent with the subgroup specificity of the disease as the envelope glycoprotein, gp70, is involved both in defining the subgroup of the virus and acting as a binding protein for target cells. Unlike leukaemia and lymphosarcoma, which have latent periods of months or years, erythroid hypoplasia develops rapidly after productive infection. In FeLV-induced immunosuppression which develops rapidly after infection, it has been shown that an isolated viral protein can produce many of the effects associated with this syndrome, including the suppression of lymphocyte mitogenesis²². Thus it is important to determine whether the induction of this disease is dependent on integration of the FeLV-C genome into target cells or whether the binding of viral proteins may suppress the activity of erythroid precursor or accessory cells.

We thank Margaret Booth and Mathew Golder for technical assistance. This work was supported by the CRC and the Leukaemia Research Fund.

Received 23 October 1981; accepted 26 January 1982.

1. Jarrett, W. F. H., Martin, W. B., Crichton, G. W., Dalton, R. G. & Stewart, M. F. *Nature* **202**, 566-567 (1964).
2. Jarrett, W., Jarrett, O., Mackey, L., Laird, H. M. & Hardy, W. D. Jr *J. natn. Cancer Inst.* **53**, 833-841 (1973).
3. Hardy, W. D. Jr, Old, L. J., Hess, P. W., Essex, M. & Cotter, S. M. *Nature* **244**, 266-169 (1973).
4. Mackey, L. J. *Vet. Rec.* **96**, 5-11 (1975).
5. Hardy, W. D. Jr *et al. Cancer Res.* **36**, 582-588 (1976).
6. Sarma, P. S. & Log, T. *Virology* **44**, 352-358 (1971).
7. Jarrett, O., Hardy, W. D. Jr, Golder, M. C. & Hay, D. *Int. J. Cancer* **21**, 334-337 (1978).
8. Mackey, L., Jarrett, W., Jarrett, O. & Laird, H. M. *J. natn. Cancer Inst.* **54**, 209-217 (1975).
9. Hardy, W. D. Jr in *Feline Leukaemia Virus* (eds Hardy, W. D. Jr, Essex, M. & McClelland, A. J.) 3-31 (Elsevier/North Holland, Amsterdam, 1980).

10. Hoover, E. A., Kociba, G. J., Hardy, W. D. Jr & Yohn, D. S. *J. natn. Cancer Inst.* **54**, 209-217 (1975).
11. Jarrett, O. & Russell, P. H. *Int. J. Cancer* **21**, 466-472 (1978).
12. Anderson, L., Wilson, R. & Hay, D. *Res. Vet. Sci.* **12**, 579-583 (1971).
13. Axelrad, A. A., McCleod, D. L., Shreeve, M. M. & Health, D. S. in *Hemopoiesis in Culture* (DHEW Publication no. 74-205 (ed. Robinson, W. A.) (NIH, 1973).
14. Pluznik, D. H. & Sachs, L. *J. cell. Physiol.* **66**, 319-324 (1965).
15. Bradley, T. R. & Metcalf, D. *Aust. J. exp. Biol. med. Sci.* **44**, 287-300 (1966).
16. Hoover, E. A. & Kociba, G. J. *J. natn. Cancer Inst.* **53**, 1277-1284 (1974).
17. Jarrett, O. in *Cold Spring Harb. Conf. Cell Proliferation* Vol. 7 (eds Essex, M., Todaro, G. & zur Hausen, H.) 603-611 (Cold Spring Harbor Laboratory, 1980).
18. Krantz, S. B. *Br. J. Haemat.* **25**, 1-6 (1973).
19. Marmont, A., Peschle, C., Sanguineti, M. & Condorelli, M. *Blood* **45**, 247-261 (1975).
20. Kagan, W. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **73**, 2890-2894 (1976).
21. Nathan, D. G. *et al. J. exp. Med.* **147**, 324-339 (1978).
22. Olsen, R. G., Mathes, C. E., Hedebrand, L. C., Hoover, E. A. & Nichols, W. S. *Am. J. Path.* **98**, 857-860 (1980).
23. Jarrett, O. *Int. J. Cancer* (submitted).
24. Guibert, L. J. & Iscove, N. N. *Nature* **263**, 594-595 (1976).
25. Tambourin, P. E., Wendling, F., Gallien-Lartigue, D. & Hualme, D. *Biomedicine* **19**, 112-116 (1973).

Two genetic loci control the murine immune response to A-gliadin, a wheat protein that activates coeliac sprue

Martin F. Kagnoff

Department of Medicine, University of California at San Diego, La Jolla, California 92093, USA

Coeliac sprue is a disease in humans that is characterized by small intestinal mucosal injury and malabsorption. Ingestion of dietary gluten by susceptible individuals results in damage to the small intestine. The gliadin fractions of wheat gluten and similar proteins in rye and barley are known to activate the disease process¹. The major histocompatibility complex (MHC) alleles *HLA-B8*, *Dw3* and *DRw3* occur at a high frequency in coeliac sprue²⁻⁵, suggesting that immunoregulatory mechanisms may be important in the pathogenesis of the disease. We recently reported that murine antibody responses to A-gliadin, a well characterized protein known to activate coeliac sprue^{1,6-9}, are regulated by genes that map to *H-2*, the murine MHC¹⁰. However, it appeared that an additional genetic locus also might be important in regulating the immune response to gliadin proteins. We show here that two separate genetic regions determine the murine antibody response to A-gliadin: *H-2* on chromosome 17 and the immunoglobulin heavy chain (*C_H*) allotype locus (*Igh*) on chromosome 12. Possible ways in which these two loci could determine the response against A-gliadin are discussed.

Our studies took advantage of *C_H* and *H-2* congenic strains of mice to assess the IgG humoral antibody response to A-gliadin. Mice bearing the *H-2^d* haplotype are high responders to A-gliadin whereas those bearing the *H-2^k*, *H-2^s*, *H-2^b* and *H-2^a* haplotypes are low responders¹⁰. Groups of age- and sex-matched mice were injected intraperitoneally (i.p.) with 25 µg A-gliadin in complete Freund's adjuvant (CFA), and bled 18, 29 and 42 days after immunization. IgG anti-A-gliadin antibody levels were measured by radioimmunoassay¹⁰.

BALB/c (*H-2^d*, *Igh^a*) and C.B20 (*H-2^d*, *Igh^b*) congenic mice have different *Igh* haplotypes but share the same *H-2* and background genes. As shown in Fig. 1, IgG anti-A-gliadin responses in C.B20 mice were significantly lower than the BALB/c controls. Similarly reduced anti-A-gliadin responses were seen in the BALB/c *Igh* congenic C.AL20 strain (*H-2^d*, *Igh^d*). The C.B20.B and C.B20 congenic strains differ only at *H-2* whereas the C.B20.B and BALB/c congenic strains differ both at *H-2* and the *C_H* locus. As shown in Fig. 1, anti-A-gliadin responses in C.B20.B mice were significantly lower than in C.B20 and BALB/c mice.

Figure 2 compares the IgG anti-A-gliadin response in selected inbred murine strains and shows that responses in BALB/c allotype congenic mice (C.AL20, C.B20) were >50% lower

than in the BALB/c strain. C.B20 ($H-2^d$, Igh^b) and B10.D2 ($H-2^d$, Igh^b) mice that share the same $H-2$ and Igh haplotypes, but differ markedly in background genes, responded similarly. Further, responses in C.B20.B ($H-2^b$, Igh^b), C57BL/6 ($H-2^b$, Igh^b) and C57BL/10 ($H-2^b$, Igh^b) mice that share the same $H-2$ and Igh haplotypes were similar despite the presence of different background genes (Fig. 2).

The BALB.B ($H-2^b$, Igh^a) congenic strain bears the $H-2^b$ low responder haplotype but has the BALB/c Igh^a haplotype and background genes. BALB.B mice produced significantly greater IgG anti-A-gliadin responses than other $H-2^b$ strains not bearing the Igh^a haplotype. In contrast, mice with the $H-2^k$ haplotype were low responders regardless of the Igh haplotype. Finally, BAB/14 mice ($H-2^d$, $Igh^{b/a}$) bear the Igh^b allotype markers of the C.B20 and C57BL strain but have some of the immunoglobulin heavy chain variable region (V_H) genes of BALB/c. BAB/14, C.B20 and BALB/c Igh congenic mice have the same $H-2$ and background genes. As shown in Fig. 2, BAB/14 anti-A-gliadin responses were significantly lower than those in BALB/c mice.

These data demonstrate that genes at or closely linked to the murine C_H locus have a major role in determining the magnitude of the IgG response to A-gliadin. Thus, we have now shown two separate genetic regions that are associated with regulation of the anti-A-gliadin response—the major histocompatibility complex ($H-2$) on chromosome 17¹⁰ and the C_H locus on chromosome 12. Background genes did not significantly influence the anti-A-gliadin response in the strain combinations tested.

Anti-A-gliadin responses varied two- to threefold between $H-2$ identical strains bearing different Igh haplotypes. However, all $H-2^d$ strains produced significant levels of IgG anti-A-gliadin antibody irrespective of Igh haplotype, and $H-2^k$ strains behaved as low responders regardless of whether or not they bore the Igh^a haplotype. These results suggest that $H-2$ -linked genes predominate in determining the response status to A-gliadin. Within $H-2$, genes mapping to the $I-A$ subregion seem to be most important in regulating specific anti-A-gliadin responses (M.F.K., unpublished data). Products coded for by $H-2$ - and C_H -linked genes may function independently or interact in controlling the anti-A-gliadin response.

Our data do not distinguish whether control of the anti-A-gliadin response is mediated by C_H genes or genes closely linked to that locus. Structural genes closely linked to the C_H locus are known to encode variable region idiotype determinants on immunoglobulin heavy chains (V_H genes)¹¹⁻¹³. Such idiotype

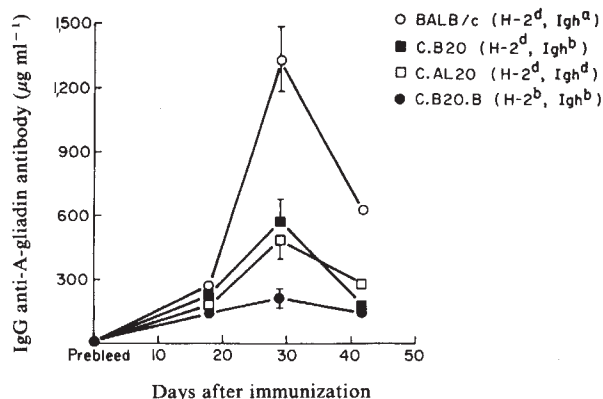


Fig. 1 IgG anti-A-gliadin responses in BALB/c allotype congenic mice. A representative experiment is shown in which groups of four to six 10-week-old female mice were immunized with 25 µg A-gliadin i.p. in CFA. Prebleed sera and sera obtained from individual mice on days 18, 29 and 42 after immunization were assayed for specific IgG anti-A-gliadin antibody by radioimmunoassay. Vertical bars represent s.e.m. On day 29, C.B20, C.AL20 and C.B20.B responses were significantly less ($P < 0.001$) than BALB/c responses, and C.B20.B responses were significantly lower than the C.B20 responses ($P < 0.01$).

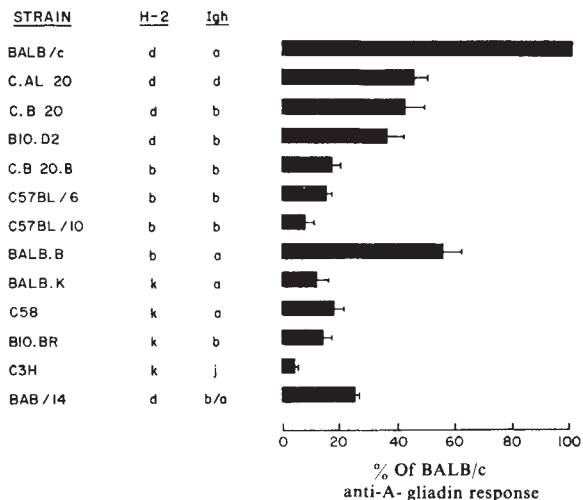


Fig. 2 Comparison of IgG anti-A-gliadin responses in inbred strains of mice. Groups of 4–12 mice of each strain were immunized with 25 µg A-gliadin i.p. in CFA. Results compare day 29 IgG anti-A-gliadin responses in various strains with the responses in age- and sex-matched BALB/c mice that were immunized and assayed in parallel. In all strains responses on day 29 were greater than those on days 18 or 42. Anti-A-gliadin antibody was not detected in prebleed sera from any of these strains. Horizontal error bars represent s.e.m. Responses in C3H/An and C3H/St mice were the same (not shown). However, IgG anti-A-gliadin responses were 10-fold greater in C3H.OH than in either C3H strain. The C3H.OH recombinant strain is congenic to C3H but bears the $H-2^d$ haplotype at K and in the I region of $H-2$.

determinants might regulate the anti-A-gliadin response through their involvement in lymphocyte receptor recognition of A-gliadin determinants or cellular cooperation. For example, it could be postulated that T-cell help for the IgG anti-A-gliadin response is directed to V_H -coded idiotype determinants¹⁴⁻¹⁸ on anti-A-gliadin antibody associated with the B cell. This interpretation predicts a unique pattern of idiotype expression to A-gliadin determinants among strains of mice bearing different immunoglobulin heavy chain allotypes. Alternatively, C_H genes would seem to control the anti-A-gliadin response if allotype-specific T-cell help or suppression^{19,20} differs among the various Igh haplotypes. However, T regulatory cells that appear to be allotype specific might be idiotype rather than allotype specific due to a close linkage and nonrandom association between C_H and V_H products²¹. Minor histocompatibility genes that are closely linked to the C_H locus²²⁻²⁴ might control the anti-A-gliadin response by coding for lymphocyte cell-surface structures that are important in the recognition or binding of A-gliadin determinants, or in cellular cooperation. Finally, C_H -linked genes postulated to code for the constant regions of T-cell receptors on helper and suppressor cells²⁵ may be important in the regulation of the anti-A-gliadin response.

To determine whether V_H genes regulate the anti-A-gliadin response, we used BAB/14 ($H-2^d$, $Igh^{b/a}$) mice. BAB/14 is an intra-V-region recombinant strain that bears the C_H genes of C.B20 and C57BL (that is, Igh^b) but expresses many of the V-region idiotype markers of the BALB/c strain^{26,27}. The finding that IgG anti-A-gliadin responses in BAB/14 mice were significantly lower than in BALB/c mice could argue against V_H structural gene control of the anti-A-gliadin response. On the other hand, it is known that the BAB/14 strain expresses some C57BL idiotypes and fails to express certain BALB/c idiotypes^{26,27}. Thus, the BAB/14 data may suggest that the V_H structural genes determining the putative BALB/c A-gliadin idiotype map closer to the C_H allotype locus than to the V_H genes that determine the BALB/c idiotype to determinants like $\alpha(1,3)$ linkages on dextran B1355^{13,28}. Alternatively, differences in the anti-A-gliadin response between BAB/14 and BALB/c mice might reflect an allotype-linked regulatory

gene that controls the phenotypic expression of the putative anti-A-gliadin idiotype.

Coeliac sprue is unique among the HLA-associated diseases in that the proteins known to activate this disease are well characterized^{1,6-9}. Although more than 80% of patients in certain populations with coeliac sprue have the *HLA-B8*, *Dw3* and *DRw3* alleles²⁻⁵, many individuals with these same HLA alleles ingest dietary gluten and do not develop disease. Further, family studies indicate that HLA-identical siblings are often disparate for the development of coeliac sprue^{5,29}. Such

observations have led to the suggestion that a second genetic region may be important in the pathogenesis of this disease³⁰. Our studies with A-gliadin in mice suggest that the second genetic locus may be linked to the genes that code for immunoglobulin heavy chain allotype. Our preliminary studies in man indicate that this is indeed the case.

This work was supported by NIH grant AM 17276 and a grant from the National Foundation for Ileitis and Colitis Inc. I thank Drs D. D. Kasarda and J. E. Bernardin for A-gliadin preparations and Mr R. Austin for technical assistance.

Received 14 December 1981; accepted 25 January 1982.

1. Kasarda, D. D. in *Food, Nutrition, and Evolution* (eds Walcher, D. & Kretchmer, M.) 201-216 (Masson, New York, 1981).
2. Faichuk, Z. M., Rogentine, G. N. & Strober, W. *J. clin. Invest.* **51**, 1602-1605 (1972).
3. Keuning, J. J. *et al. Lancet* **i**, 506-507 (1976).
4. Ek, J., Albrechtsen, D., Solheim, B. G. & Thorsby, E. *Scand. J. Gastroenterol.* **13**, 229-233 (1978).
5. Polanco, I. *et al. in The Genetics of Coeliac Disease* (ed. McConnell, R. B.) 211-230 (MTP Press, Lancaster, 1981).
6. Platt, S. G. & Kasarda, D. D. *Biochim. biophys. Acta* **243**, 407-415 (1971).
7. Kasarda, D. D., Bernardin, J. E. & Nimmo, C. C. in *Advances in Cereal Science and Technology* Vol. 1 (ed. Pomeroy, Y.) 158-236 (American Association of Cereal Chemists, St Paul, Minnesota, 1976).
8. Kasarda, D. D. *Ann. Technol. Agric.* **29**, 151-173 (1980).
9. Autran, J.-C., Lew, E. J.-L., Nimmo, C. C. & Kasarda, D. D. *Nature* **282**, 527-529 (1979).
10. Trefts, P. E. & Kagnoff, M. F. *J. Immun.* **126**, 2249-2252 (1981).
11. Lieberman, R., Potter, M., Mushinski, E. B., Humphrey, W. Jr & Rudikoff, S. *J. exp. Med.* **139**, 983-1001 (1974).
12. Riblet, R. *et al. Eur. J. Immun.* **5**, 775-777 (1975).
13. Lieberman, R. *Springer Semin. Immunopath.* **1**, 7-30 (1978).
14. Bottomly, K. & Mosier, D. E. *J. exp. Med.* **150**, 1399-1409 (1979).
15. Woodland, R. & Cantor, H. *Eur. J. Immun.* **8**, 600-606 (1978).
16. Hetzelberger, D. & Eichmann, K. *Eur. J. Immun.* **8**, 846-852 (1978).
17. Adorini, L., Harvey, M. & Sercarz, E. E. *Eur. J. Immun.* **9**, 906-909 (1979).
18. Trefts, P. E., Rivier, D. A. & Kagnoff, M. F. *Nature* **292**, 163-165 (1981).
19. Herzenberg, L. A. *et al. J. exp. Med.* **144**, 330-344 (1976).
20. Jacobson, E. B. *J. exp. Med.* **148**, 607-612 (1978).
21. Slack, J., Der-Balian, G. P., Nahm, M. & Davie, J. M. *J. exp. Med.* **151**, 853-862 (1980).
22. Rolink, T., Eichmann, K. & Simon, M. M. *Immunogenetics* **7**, 321-336 (1978).
23. Riblet, R. & Congleton, C. *Immunogenetics* **5**, 511 (1977).
24. Owen, F. L., Finnegan, A., Gates, E. R. & Gottlieb, P. D. *Eur. J. Immun.* **9**, 948-955 (1979).
25. Owen, F. L., Riblet, R. & Taylor, B. A. *J. exp. Med.* **153**, 801-810 (1981).
26. Pisetsky, D. S., Riordan, S. E. & Sachs, D. H. *J. Immun.* **122**, 842-846 (1979).
27. Clafin, J. L., Wolfe, J. & Ruppert, V. J. *J. Immun.* **123**, 2088-2090 (1979).
28. Weigert, M. & Riblet, R. *Springer Semin. Immunopath.* **1**, 133-169 (1978).
29. van Rood, J. J., van Hooff, J. P. & Keuning, J. J. *Transplant Rev.* **22**, 75-105 (1975).
30. PENA, A. S. *et al. Gastroenterology* **75**, 230-235 (1978).

Radical scavenging and stimulation of prostaglandin synthesis not anti-inflammatory?

T. G. Payne, B. Dewald, H. Siegl, H. U. Gubler, H. Ott & M. Baggiolini

Preclinical Research, Sandoz Ltd, CH-4002 Basle, Switzerland

Prostaglandins (PGs) are produced from arachidonic acid by a cascade of enzymes, the initial step being the conversion of arachidonic acid to PGG₂ by cyclooxygenase. In the next step, PGG₂ is converted to PGH₂, and it is at this stage that an oxidative radical is generated, presumably $\cdot\text{OH}$, which is capable of inactivating several enzymes of the arachidonic acid cascade^{1,2}. This inactivation can be prevented by radical scavengers. MK 447, a phenolic compound with such properties, was indeed found to facilitate the conversion of PGG₂ to PGH₂ and therefore to enhance PGE₂ synthesis^{2,3}. MK447 also inhibits irritant-induced inflammation in mice and rats; this was thought to depend on radical scavenging, which led Kuehl and colleagues to suggest that PGG₂ and $\cdot\text{OH}$ are mediators of inflammation^{2,3}. We have tested MK447 analogues containing asymmetric centres and report here that although both antipodes were equipotent radical scavengers and stimulators of prostaglandin synthesis, only the (+) antipodes had anti-inflammatory and diuretic⁴ properties. This indicates that the radical scavenging properties and pharmacological actions of MK 447 and derivatives are not interdependent and questions the pro-inflammatory role of PGG₂.

Several analogues of MK 447 were prepared and tested in a cell-free system for enhancement of prostaglandin synthesis from arachidonic acid, and in the rat for inhibition of irritant-induced paw oedema and for diuretic activity. We present here data for MK 447 and two analogues, A and B, which contain an asymmetric carbon atom (Fig. 1). Compounds A and B were resolved into their optical antipodes by crystallization of the dibenzoyl tartaric acid and tartaric acid salts, respectively, the details of which will be reported elsewhere (H.O., in preparation). The compounds were tested both as racemates and as the pure isomers.

We tested the effect of the aminophenols on the formation of PGE₂ from arachidonate by a microsomal preparation from bovine seminal vesicles. In both the presence and absence of

Table 1 Effects of MK 447 and related phenolic compounds on PGE₂ formation by bovine seminal vesicle cyclooxygenase

Compound	Conc. (μM)	Relative amounts of PGE ₂ formed	
		- GSH	+1 mM GSH
None	-	100	100
MK 447	10	159	149
	100	323	211
A (racemic)	10	158	148
	100	405	236
A (+) isomer	10	170	152
	100	404	245
A (-) isomer	10	158	152
	100	412	223
B (racemic)	10	187	143
	100	391	229
B (+) isomer	10	168	143
	100	362	228
B (-) isomer	10	184	144
	100	388	241

Bovine seminal vesicle microsomes (1.85 mg protein), prepared by the method of Takeguchi *et al.*⁹, were incubated with [1-¹⁴C]arachidonic acid (33 nmol, corresponding to 0.04 μCi) in 1 ml of 0.1 M sodium phosphate buffer pH 7.4, in the presence or absence of GSH and the indicated concentrations of test substances. After 30 min at 37 °C, the reaction was terminated by acidification to pH 4. Products and non-metabolized arachidonic acid were extracted into ethyl acetate, separated by TLC (benzene/dioxane/acetic acid = 50:50:2), eluted and measured by scintillation counting. The data are presented as percentage of controls. In control assays, PGE₂ production was 217 and 1,182 ng per mg of protein per 30 min, in the absence and presence of 1 mM GSH, respectively.

reduced glutathione (GSH), all compounds stimulated strongly the formation of PGE₂ (Table 1) and PGF₂ (not shown) in a concentration-dependent manner. However, in the presence of GSH, which by itself stimulates PGE₂ formation⁵, the enhancement by the phenolic compounds was reduced. Independent of their structure and chirality, all compounds tested showed similar effects. The same results were obtained for various other phenolic compounds (data not shown). In agreement with Kuehl *et al.*³, the structural determinants essential for activity were found to be the phenolic hydroxyl group and an adjacent aminomethyl substituent.

The anti-oedematous and diuretic effects of MK 447 and compounds A and B are summarized in Table 2. MK 447 and

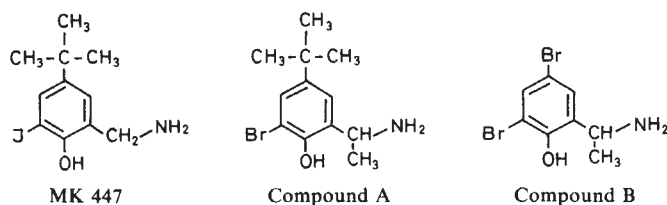


Fig. 1 Structure of MK 447 and the two optically active derivatives, compounds A and B.

both racemates of the two analogues showed moderate to strong inhibition of the inflammatory response induced by intraplantar injection of carrageenin, and markedly enhanced water and sodium excretion. Despite their close structural relationship and almost identical stimulation of prostaglandin synthesis, compounds A and B were only about one third and one tenth, respectively, as active as MK 447 in inhibiting paw oedema and enhancing diuresis. Most remarkably, both *in vivo* activities appeared to be due to the (+) isomer. As shown in Fig. 2, the (–) isomers are both much less effective than the (+) isomers or the racemates.

Comparison of the pharmacological activities of the two isomers is only valid if their absorption and fate *in vivo* are similar. This was tested in the case of compound B. The (+) and (–) antipodes were administered orally to separate groups of rats at a dose of 10 mg per kg, and urine was collected during the following 7 hours. In analogy with MK 447 (ref. 6), *O*-sulphate was expected to be the major urine metabolite. The urine was desalted using Sep-pak cartridges (Waters Associates, Massachusetts) and incubated at 55 °C for 2 h with *Helix pomatia* extract (100,000 Fishman units of β -glucuronidase and 10⁶ Roi units of arylsulphatase). The samples were then evaporated to dryness, taken up in a small volume of ethanol

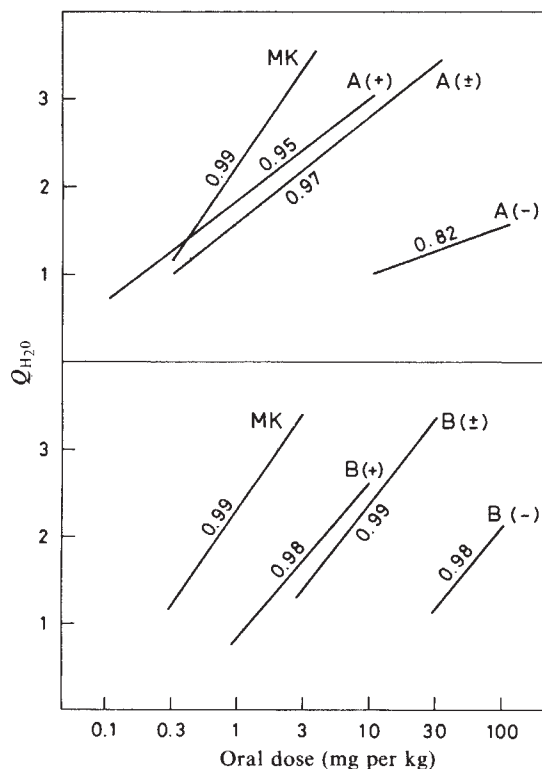


Fig. 2 Dose-dependent diuretic effect of MK 447 and related phenolic compounds in the rat (for method see Table 2), expressed as the ratio of water excretion in the treated compared with that in control animals, Q_{H_2O} . The curves were calculated by regression analysis of the values obtained for individual animals; the correlation coefficient, r , is shown for each curve. Each compound was tested at 3–5 doses in 3–4 animals per dose. MK, MK 447.

Table 2 Anti-oedematous and diuretic activities of MK 447 and related phenolic compounds in the rat

Compound	Inhibition of carrageenin- and naphthoylheparamine-induced paw oedema (ED ₅₀ ; mg per kg p.o.)		Diuretic effect in the rat (ED ₁₀₀ ; mg per kg p.o.)	
	C	NH	Water excretion	Sodium excretion
MK 447	3.9	2.8	0.7	0.75
A (racemic)	10.6	13.2	2.0	1.5
A (+) isomer	6.5	8.2	1.1	1.1
A (–) isomer	>30.0	>30.0	>100.0	>100.0
B (racemic)	12.2	14.2	6.5	7.3
B (+) isomer	4.3	4.0	4.6	4.5
B (–) isomer	>30.0	>30.0	85.0	>100.0

The carrageenin (C) oedema test¹⁰ and the naphthoylheparamine (NH) oedema test¹¹ were performed in the rat according to established methods. In both cases, paw swelling was assessed by conductometry as described by Kemper and Ameln¹². The compounds were given orally (p.o.) as a suspension in saline containing 0.5% tragacanth, 1 h before local injection of the irritant. The ED₅₀ (the average dose giving 50% inhibition of irritant-induced swelling compared with untreated controls) was determined from dose-response curves (3–4 doses per compound, 5 rats per dose). The inhibition of swelling after a dose of 30 mg per kg of the (–) isomers was 15 and 2% for compound A and 19 and 7% for compound B. Diuretic activity was assessed by the method of Flückiger *et al.*¹³. Fasted rats (200 g) were hydrated (30 ml water per kg, orally) and 2 h later, after emptying the bladder, they were given the test compounds in 50 ml of isotonic saline per kg. The total urine produced during the following 3 h was measured and tested for sodium, potassium and chloride content. Average water and sodium excretion per 200 g body weight was calculated (3–4 rats per dose) and the ED₁₀₀, the average dose inducing a 100% increase in both values, determined for 3–5 different dosages.

and analysed by TLC. Three solvent systems were used and the compounds were detected under UV light and with iodine vapours. Similar amounts of the aminophenol were obtained after administration of either antipode, indicating that they were absorbed and excreted to the same extent. In both cases, no aminophenol was detectable in urine samples which were not treated with the hydrolases, suggesting that the metabolism of the two antipodes was similar.

Our results show that the anti-oedematous effects of phenols like MK 447 and its analogues are unlikely to depend on the ability of the compounds to scavenge oxygen radicals and hence to increase the conversion of PGG₂ to PGH₂ and further metabolites. The two optical isomers of compounds A and B enhance prostaglandin synthesis to the same extent. As indicated by the experiments with compound B, both isomers have similar absorption and fate in rats. Nevertheless, they differ widely in their *in vivo* activities, the (+) isomers only being biologically comparable to MK 447. These results suggest that the theory that radicals which arise from the peroxidase-dependent reduction of PGG₂ are pro-inflammatory and that PGG₂ is a key element in the pathogenesis of inflammation^{2,3,7} needs revision.

Our results also question the notion that the main action of MK 447, that is, diuresis and saluresis⁴, is mediated by its enhancing prostaglandin synthesis. This was suggested by the observation that MK 447-induced diuresis was inhibited by indomethacin^{4,8}. However, as the (–) isomers of compounds A and B also have some diuretic activity, it cannot be ruled out entirely that an increase of prostaglandin production may contribute to the overall effect.

Received 8 December 1981; accepted 21 January 1982.

- Kuehl, F. A. Jr & Egan, R. W. *Science* **210**, 978–984 (1980).
- Kuehl, F. A. Jr, Humes, J. L., Torchiana, M. L., Ham, E. A. & Egan, R. W. in *Advances in Inflammation Research* Vol. 1 (eds Weissmann, G., Samuelsson, B. & Paoletti, R.) 419–430 (Raven, New York, 1979).
- Kuehl, F. A. Jr *et al.* *Nature* **265**, 170–173 (1977).
- Scrabine, A. *et al.* *J. Pharmac. exp. Ther.* **208**, 148–154 (1979).
- Ogino, N., Miyamoto, T., Yamamoto, S. & Hayaishi, O. *J. biol. Chem.* **252**, 890–895 (1977).
- Tocco, D. J. *et al.* *Pharmacologist* **20**, 214 (1978).

7. Egan, R. W., Galen, P. H., Beveridge, G. C., Phillips, G. B. & Marnett, L. J. *Prostaglandins* **16**, 861–869 (1978).
8. Mason, R. T. & Staszewska-Barczak, J. *J. clin. exp. Pharmac. Physiol.* **6**, 678–685 (1979).
9. Takeguchi, C., Kohno, E. & Sih, C. J. *Biochemistry* **10**, 2372–2376 (1971).
10. Winter, C. A., Risley, E. A. & Nuss, G. W. *Proc. Soc. exp. Biol. Med.* **111**, 554–557 (1962).
11. Branceni, D., Azadian-Boulanger, G. & Jequier, R. *Archs int. Pharmacodyn.* **152**, 15–24 (1964).
12. Kemper, F. & Ameln, G. *Z. ges. exp. Med.* **131**, 407–415 (1959).
13. Flückiger, E., Schlach, W. & Taeschler, M. *Schweiz. med. Wschr.* **93**, 1232–1237 (1963).

A model of the nerve impulse using two first-order differential equations

J. L. Hindmarsh & R. M. Rose

Department of Applied Mathematics and Astronomy and
Department of Physiology, University College, Cardiff,
Cardiff CF1 1XL, UK

The Hodgkin–Huxley model¹ of the nerve impulse consists of four coupled nonlinear differential equations, six functions and seven constants. Because of the complexity of these equations and the necessity for numerical solution, it is difficult to use them in simulations of interactions in small neural networks. Thus, it would be useful to have a second-order differential equation which predicted correctly properties such as the frequency–current relationship. Fitzhugh² introduced a second-order model of the nerve impulse, but his equations predict an action potential duration which is similar to the inter-spike interval³ and they do not give a reasonable frequency–current relationship. To develop a second-order model having few parameters but which does not have these disadvantages, we have generalized the second-order Fitzhugh equations², and based the form of the functions in the new equations on voltage-clamp data obtained from a snail neurone. We report here an unexpected property of the resulting equations—the x and y null clines in the phase plane lie close together when the phase point is on the recovery side of the phase plane. The resulting slow movement along the phase path gives a long inter-spike interval, a property not shown clearly by previous models^{2,4}. The model also predicts the linearity of the frequency–current relationship, and may be useful for studying detailed interactions in networks containing small numbers of neurones.

Although our model may be seen as a generalization of the Fitzhugh equations², it can be developed from first principles if it is assumed that the rate of change of membrane potential depends linearly on z (the current passed through the electrode), and y (an intrinsic current), and depends nonlinearly on membrane potential, giving

$$\dot{x} = -a(f(x) - y - z) \quad (1)$$

We assume also that the rate of change of intrinsic current is given by

$$\dot{y} = b(g(x) - y) \quad (2)$$

In these equations, a and b are constants, and equation (2) has a form which ensures that the time course of z in voltage-clamp conditions is exponential (see equation (4)). This is a simplification which ignores the voltage-dependence of the time constants.

To determine the form of $f(x)$ and $g(x)$ in equations (1) and (2), a large cell from the visceral ganglion of the pond snail, *Lymnaea stagnalis*, was clamped to a range of different voltages (x_p), and the initial ($z_{xp}(0)$) and steady-state ($z_{xp}(\infty)$) values of the clamping current were measured, giving a conventional⁵ current–voltage plot (Fig. 1). As the cell fires repetitively in the resting state, the membrane potential was biased initially by adjusting z so that the cell just stopped firing. This corresponds to an artificial stable equilibrium point, and will be used as the origin ($x = 0, y = 0$) in the phase plane. In the current–voltage plot (Fig. 1), the scale is also positioned so that the current required to maintain the equilibrium voltage is at the origin

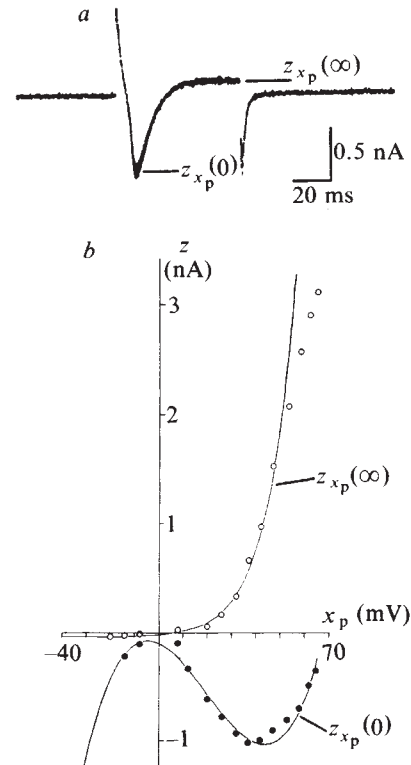


Fig. 1 Voltage-clamp data on which the model is based. In these experiments, the brain of the pond snail *L. stagnalis* was pinned out in a chamber with a sylgard base, and the outer sheath softened using pronase enzyme for a few minutes before recording. The bathing solution had the following composition (mM): Na^+ , 24.0; K^+ , 2.0; Ca^{2+} , 4.0; Mg^{2+} , 2.0; Cl^- , 38.0; $\text{H}_2\text{PO}_4^{2-}$, 0.1; Na-HEPES, 35.4. A large cell ($\sim 200 \mu\text{m}$) in the visceral ganglion was impaled with a bevelled glass microelectrode filled with 3 M KCl, and having a resistance of 4–5 M Ω . The cell was clamped using the single-electrode voltage-clamp technique of Wilson and Goldner⁷. *a*, Current change recorded for a voltage step (x_p) of 25 mV from the equilibrium point (cell held at an actual potential of -60 mV, which is the equilibrium point, $x = 0$, in the measurements). The initial ($z_{xp}(0)$, or peak inward) and final ($z_{xp}(\infty)$, or steady-state outward) currents are indicated, the final currents being measured after 60 ms in this case. *b*, Initial and final currents obtained from records such as shown in *a* are plotted for different values of clamp potential (x_p) measured from the equilibrium point. Negative values of z are inward, and positive values are outward currents. Final currents were measured 100 ms after onset of voltage step. Different animal than in *a*, but curves are typical for this cell (experiment repeated five times). The $z_{xp}(0)$ points have been fitted by the least-squares method, using the cubic equation $f(x) = cx^3 + dx^2 + ex + h$, where $c = 0.00017$, $d = 0.001$, $e = 0.01$ and $h = 0.1$. The $z_{xp}(\infty)$ points were fitted by an exponential of the form $g(x) = qe^{rx} - s$, where $q = 0.024$, $r = 0.088$ and $s = 0.046$.

($z = 0$). Therefore, at the equilibrium position $x = y = \dot{x} = \dot{y} = z = 0$. The time course of the current, after the onset of a step in voltage to x_p , is given (from equation (1) with $\dot{x} = 0$) by

$$z_{xp}(t) = f(x_p) - y(t) \quad (3)$$

where y satisfies equation (2) with $x = x_p$.

Therefore, $z_{xp}(t)$ is given by

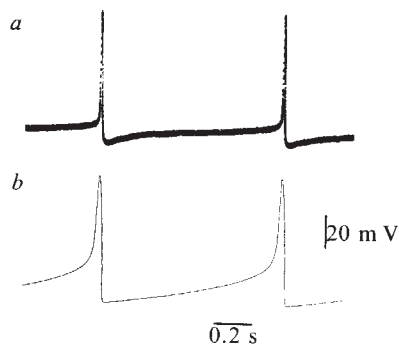
$$z_{xp}(t) = f(x_p) - g(x_p)(1 - e^{-bt}) \quad (4)$$

and hence

$$z_{xp}(0) = f(x_p), \text{ and } z_{xp}(\infty) = f(x_p) - g(x_p) \quad (5)$$

To obtain $f(x_p)$, the values of the initial currents ($z_{xp}(0)$; Fig. 1*b*) were fitted by a cubic equation. To obtain $g(x_p)$, an exponential function was fitted to the value of the final currents ($z_{xp}(\infty)$; Fig. 1*b*), and $g(x_p)$ was obtained from equation (5). The time constant b was estimated from voltage-clamp current records

Fig. 2 Comparison between *a*, actual and *b*, calculated waveforms for the same cell as in Fig. 1*b*. The calculated waveform is the solution to equations (7) and (8) for a current of $z = 0.008$, with the constants $a = 5,400 \text{ mV s}^{-1}$, nA^{-1} , $b = 30 \text{ s}^{-1}$ and using the functions $f(x)$ and $g(x)$ shown in Fig. 1*b*. Numerical solutions were obtained using the Runge-Kutta-Merson method⁸ with variable step length. In these calculations we found that repetitive firing occurred down to $z = -0.026$ rather than $z = 0$ as expected (procedure outlined in the text). This was due to the approximations adopted, which slightly changed the values of $f(x)$ and $g(x)$ at the origin. As the computed equilibrium point is at $z = 0.034$ ($=0.008 \pm 0.026$) in Fig. 4. The recorded action potentials were obtained in current-clamp conditions with an experimental z value of 0.04 and were photographed ~ 1 min after imposing the current step.



(Fig. 1*a*), which were assumed to follow an exponential time course, and the time constant a was obtained by applying a current step to the system at rest in the equilibrium position (0, 0). If the step in current is $z(t) = z_p$ for $t \geq 0$ then a may be obtained from the slope of the voltage change immediately after the onset of the current-clamp step. This is given by equation (1) with $f(0) = 0$ and $y(0) = 0$:

$$\dot{x}(0) = az_p \quad (6)$$

Thus the assumed form for our equations is

$$\dot{x} = -a(f(x) - y - z) \quad (7)$$

$$\dot{y} = b(f(x) - q e^{rx} + s - y) \quad (8)$$

where $f(x) = cx^3 + dx^2 + ex + h$, and a, h, q, r and s are constants.

After measuring a and b , and fitting cubic and exponential functions to the $z_{sp}(0)$ and $z_{sp}(\infty)$ data of Fig. 1*b*, the solutions of equations (7) and (8) were obtained by numerical integration.

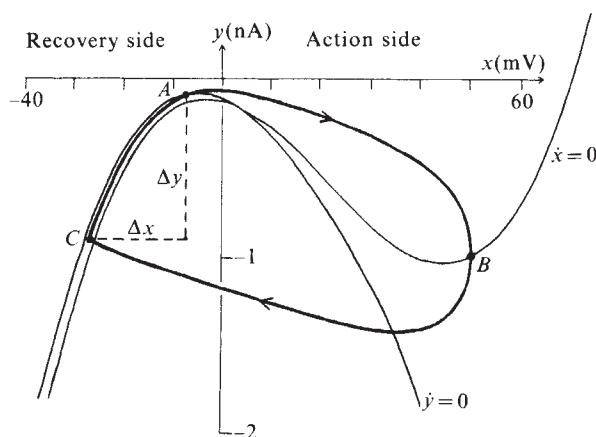


Fig. 3 Phase plane representation of the limit cycle solution to equations (7) and (8). The values of the constants are the same as in Fig. 2, except that a z value of 0.033 has been used so that the equilibrium point can be more clearly distinguished. The $\dot{x} = 0$ curve has been displaced horizontally to the right by 2–3 mV so that the separation between the x and y null clines on the recovery side of the diagram can be seen more easily. The $\dot{x} = 0$ curve is similar to that of Fitzhugh², but the $\dot{y} = 0$ curve bends round to follow the $\dot{x} = 0$ curve in the lower left-hand quadrant, which is significantly different from the Fitzhugh model. The limit cycle is divided into an action potential phase ($A \rightarrow B \rightarrow C$) and a recovery phase ($C \rightarrow A$), during which the phase point moves slowly between the x and y null clines. The differences Δx and Δy are, respectively, the changes in membrane potential and intrinsic current during recovery.

Figure 2 compares the calculated $x(t)$ waveform with that recorded from the cell. The model clearly predicts a ratio of spike duration to inter-spike interval which is similar to that actually recorded, although the calculated action potential is of rather longer duration than the recorded spike; the pacemaker potential also rises to a higher threshold in the model. Some of these differences may be corrected by taking into account the voltage dependency of the time constants, but these modifications are not introduced here because they detract from the simplicity of the model.

Although this model is not as accurate as more complicated models of repetitive firing⁶, it has a conceptual advantage which becomes apparent when the oscillation is plotted onto the phase plane. The main difference between this model and that of Fitzhugh² is the introduction of $g(x)$ into the equation for $\dot{y}$. By using the phase diagram (Fig. 3) it becomes easy to understand why this modification is so important in lengthening the recovery phase between successive action potentials. In the phase plane, when the limit cycle crosses the $\dot{x} = 0$ curve at C (Fig. 3), it is constrained to move along the narrow path between the $\dot{x} = 0$ and $\dot{y} = 0$ curves. Because the phase path is so close to both the x and y null clines, progress along the limit cycle will be much slower from C to A (the pacemaker or recovery phase) than from A through B to C (the action potential part of the cycle). By contrast, in Fitzhugh's model the $\dot{x} = 0$ and $\dot{y} = 0$ curves diverge markedly on both sides of the diagram, so that x and y change almost as rapidly during the pacemaker phase as during the spike itself.

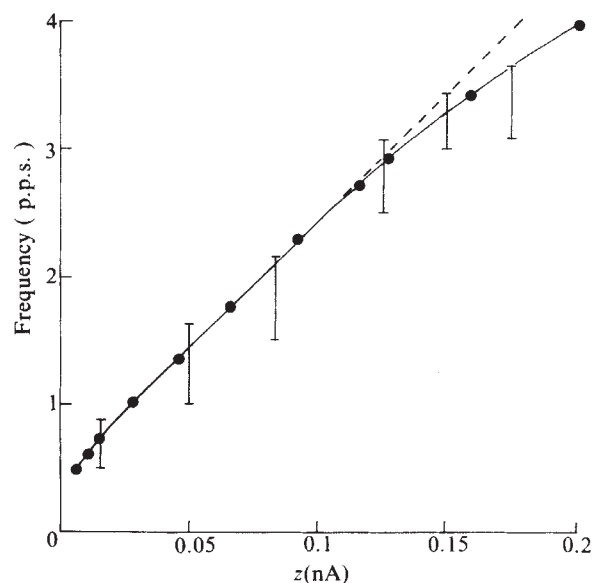


Fig. 4 Comparison of the calculated and experimental frequency-current relationship (pulses per second). The calculated frequencies (●) were derived from numerical integration of equations (7) and (8) using the same parameter values as in Fig. 2 but with different values of z . Calculated frequencies are plotted relative to the computed equilibrium point ($z = -0.026$; see Fig. 2). Experimental z values were measured relative to the experimental equilibrium point (at which $z = 0$; there were no action potentials). Over most of the range there is a linear relationship between frequency and applied current, with both experimental and calculated curves having a similar slope. At higher values of z , the calculated curve and the experimental points begin to approach a limiting frequency at a similar z value, although the experimental points appear to fall off more sharply from linearity (broken line) than the theoretical prediction. This particular cell also shows adaptation of firing frequency to an applied current step, so that the values for the experimental frequencies lie within the upper (=initial frequency in response to step) and lower (=final frequency) limits of the bar lines. Thus a precise comparison cannot be made because the model has not yet been modified to include adaptation. Experimental points were obtained from a different animal than that used in Fig. 1*b*, but the experiment was repeated four times and gave consistent results.

Numerical solution may also be used to predict the frequency-current relationship. The computed values of the frequency are linearly related to the current (Fig. 4) and compare reasonably well with the experimental values. It is also possible to derive an expression for the frequency-current relationship fairly simply. From equations (1) and (2) we see that

$$bx + ay = ab(g(x) - f(x)) + abz \quad (9)$$

Then the expression $g(x) - f(x)$ ($= -z(\infty)$; equation (5)) may be approximated for $x < 0$ by a small positive constant M (see $z_x(\infty)$ in Fig. 1b). Integrating (9) we obtain

$$b\Delta x + a\Delta y = abMT + abzT \quad (10)$$

where Δx and Δy are the changes in the values of x and y during recovery (from C to A in Fig. 3), which occurs in time T , say. If we ignore the spike duration, the frequency f of oscillation is given by

$$f = \frac{1}{T} = \frac{ab(z + M)}{b\Delta x + a\Delta y} \quad (11)$$

from which we can see that, to the extent that Δx and Δy are independent of z , f is linearly related to z (see Fig. 4).

The model therefore predicts the frequency-current relationship as well as giving a reasonably accurate representation of the time course of the membrane potential change. It also provides a simple explanation of why the recovery time is much longer than the spike duration. Because these equations are simpler to understand than the four-dimensional Hodgkin-Huxley equations¹, and can be solved numerically in a shorter time, they may be useful in studying small neural networks.

We thank Dr W. A. Wilson for supplying us with further information on the single electrode voltage-clamp, and Mr W. L. Barry and Mr T. Ford for constructing the voltage-clamp apparatus.

Received 1 September 1981; accepted 26 January 1982.

- Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **117**, 500-554 (1952).
- Fitzhugh, R. *Biophys. J.* **1**, 445-466 (1961).
- Jack, J. J. B., Noble, D. & Tsien, R. W. (eds) *Electric Current Flow in Excitable Cells*, 330 (Clarendon, Oxford, 1975).
- Kokoz, Yu. M. & Krinskii, V. J. *Biophysika* **18**, 878-885 (1973).
- Frank, K. & Tauc, L. in *Cellular Function of Membrane Transport* (ed. Hoffman, J.) 113-134 (Prentice-Hall, New Jersey, 1964).
- Connor, J. A. & Stevens, C. F. *J. Physiol., Lond.* **213**, 31-53 (1971).
- Wilson, W. A. & Goldner, M. M. *J. Neurobiol.* **6**, 411-422 (1975).
- Merson, R. H. in *Proc. of Symp. on Data Processing* (Weapons Research Establishment, South Australia, 1957).

Induction of helical liposomes by Ca^{2+} -mediated intermembrane binding

Ke-Chun Lin*, Robert M. Weis & Harden M. McConnell

Stauffer Laboratory for Physical Chemistry, Stanford University, Stanford, California 94305, USA

Extensive freeze-fracture electron microscopic studies of Ca^{2+} -containing mixtures of an acidic phospholipid (cardiolipin) and phosphatidylcholines¹⁻⁵, have revealed certain replica features which have been interpreted as regions of membrane-membrane attachment³⁻⁵. Spin-label paramagnetic resonance spectra of mixtures of another acidic phospholipid (phosphatidylserine) and phosphatidylcholines have provided evidence for Ca^{2+} -mediated lateral phase separations of the acidic phospholipids⁶⁻⁸, and their involvement in intermembrane binding⁷. We report here that in the presence of Ca^{2+} , binary mixtures of cardiolipin and phosphatidylcholines readily form tightly folded single- and double-helical liposomes; we invoke membrane-membrane attachment to explain this phenomenon.

* Permanent address: Department of Biophysics, Beijing Medical College, Beijing, China.

Ethanol solutions of bovine heart cardiolipin (Ca^{2+} -free; Sigma) and dimyristoylphosphatidylcholine (Sigma) were evaporated to dryness to form binary mixtures containing 37 mol % cardiolipin (1.7 mg total lipid). The lipid was dissolved in chloroform/methanol (10:1 v/v) and evaporated to dryness, then 160 μl phosphate-buffered saline (PBS) were added with or without 2 mM EDTA and the mixture was incubated at 45 °C for 5 min, vortexed for 10-15 s, then diluted 20-fold with PBS. All experiments were carried out at 25-26 °C, well above the chain-melting transition of dimyristoylphosphatidylcholine (23 °C).

We used two experimental protocols: (1) after addition of PBS+EDTA and vortexing, the dilution with PBS contained no Ca^{2+} . The sample was placed on a microscope slide with an elevated coverslip (100 μm) and a concentrated CaCl_2 solution (0.01-1 M) was added from one side. Ca^{2+} ions gradually diffused through the solution, allowing observation of helix formation (see Fig. 1a-d). (2) For the mixture incubated without EDTA, the subsequent dilution with PBS contained various low concentrations of Ca^{2+} (10^{-6} - 10^{-3} M).

Single and double helix formation, observed using protocol (1), is shown in Fig. 1. Long tubes were often observed in the absence of Ca^{2+} , together with liposomes having typical irregular shapes. The presence of regular tubes and irregular liposomes may be explained by possible lipid compound formation and phase separations in binary mixtures of these lipids⁹. The tubes were almost perfect cylinders, although end defects were frequently seen, and were usually multilamellar but occasionally unilamellar, as judged by microscopic observation of pronounced brownian motion, characteristic of unilamellar liposomes¹⁰. Photobleaching recovery experiments measuring diffusion parallel to the long axis of the tubes yielded lateral diffusion coefficients typical of fluorescent lipid probes ($D \sim 10^{-8} \text{ cm}^2 \text{ s}^{-1}$).

Protocol (1) provided evidence that the rate of helix formation increased with increasing Ca^{2+} concentration. Protocol (2) permitted the determination of the minimum concentration of Ca^{2+} ion required for helix formation; compared with controls (no Ca^{2+}), significantly more helices were formed at Ca^{2+} concentrations as low as 10^{-6} M.

The thermodynamic driving force for helix formation is the Ca^{2+} -mediated membrane-membrane binding. At higher Ca^{2+} concentrations, helices rapidly collapse on themselves forming a variety of more complex structures. It is possible that helices represent kinetic intermediates in the Ca^{2+} -mediated transformation of tubes to more complex structures, as each step of helix formation represents an increase in the area of membrane-membrane contact. The formation of a typical single helix is simple; it is initiated at one end, and is smooth and continuous. However, double-helix formation is sometimes complex; it initiates from a hairpin loop and progresses in a stepwise manner. A complete regular double helix with no loose ends can initiate at a hairpin loop but not in the centre of a tube; this requires sliding as well as twisting of one tube relative to the other (see Fig. 1a-c).

The rate at which helices form in the presence of Ca^{2+} is consistent with earlier proposals concerning Ca^{2+} -mediated lateral phase separations of acidic phospholipids⁶⁻⁸, and with the rates of lateral diffusion of these lipid molecules reported here. Ca^{2+} chelation by phosphate groups should be favoured when these groups have a three-dimensional configuration, rather than a two-dimensional configuration provided by a single membrane.

The forces opposing helix formation are weak. The shear modulus of a fluid membrane is essentially zero; a negligible force is required to twist one region of a lipid tube relative to another. We observed no distortion of the circularity of the tubes, thus there was no detectable volume or surface area change of the tubes on helix formation. Note that the outer diameter of the double helix in Fig. 1 is twice the diameter of the tube. No stretch or compression elastic energy is required to convert a cylindrical tube into a helix because lipids on the

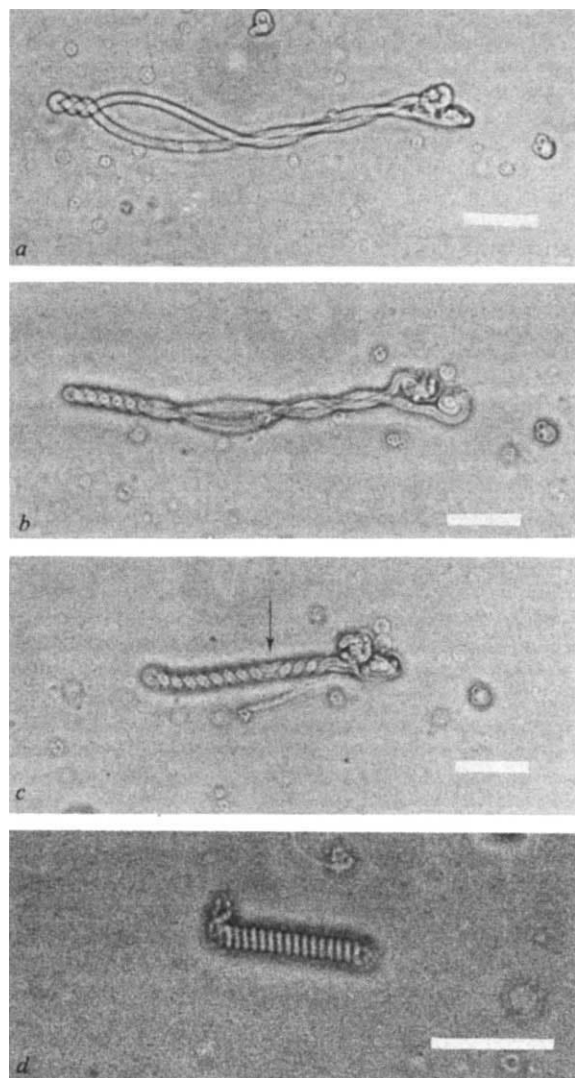


Fig. 1 Helical liposomes derived from binary mixtures containing 37 mol % cardiolipin and 63 mol % dimyristoylphosphatidylcholine as Ca^{2+} ions diffuse from a 0.01 M CaCl_2 solution at the edge of a coverslip into the region where the liposomes are located. *a-c* Illustrate the formation of a double helix initiated on the left at a hairpin loop, and on the right in an extended region of membrane-membrane contact. In *a* and *b* both helices are right-handed. In subsequent photographs (not shown) the helix at the right-hand end became unwound as the helix on the left continued to form. Subsequently, as the right-handed helix to the left continued to form, a left-handed helix formed on the right, and the two helices met in a soliton-like defect indicated by the arrow in *c*. A single helix is shown in *d*. Scale bars, 25 μm . The elapsed time from *a* to *c* is ~ 5 min; double-helix formation is not continuous but proceeds in a stepwise fashion. A similar length of time was observed for the continuous formation of the single helix shown completed in *d*.

inside of the helix can flow to the outside on bending. Helfrich¹¹ has estimated the curvature elastic energy for bending a lipid tube (of radius r) so that it has a radius of curvature R : to first-order approximation this is $\pi K r / 2 R^2$ per unit length, where K is the curvature elastic modulus. For egg lecithin, $K \approx 23 \times 10^{-12}$ erg (ref. 11). Assuming one lipid molecule per 10 Å along a contact line between helical tubes, and using the above values of K , we calculate the curvature elastic energy of lipid touching this line to be $0.013 n \text{ kcal mol}^{-1}$ for a tube with radius $r = 1 \mu\text{m}$ and having n bilayer membranes. Thermal energy, RT , at room temperature is $0.6 \text{ kcal mol}^{-1}$. Clearly, weak membrane-membrane binding energies can overcome curvature elastic energy and stabilize helices.

We acknowledge that Dr Alec Bangham has independently observed the formation of tubular and helical liposomes on

hydration of egg lecithin with saline solution, and we thank him for supplying photographs of these helical liposomes. The double helices of egg lecithin are very similar to those reported here. This work was supported by NSF grant PCM 8021993.

Received 23 November 1981; accepted 21 January 1982.

1. Verkleij, A. J., Mombers, C., Leunissen-Bijvelt, J. & Ververgaert, P. J. J. Th. *Nature* **279**, 162-163 (1979).
2. Verkleij, A. J., Mombers, C., Gerritsen, W. J., Leunissen-Bijvelt, J. & Cullis, P. R. *Biochim. biophys. Acta* **555**, 358-361 (1979).
3. Miller, R. G. *Nature* **287**, 166-167 (1980).
4. Hui, S. W. & Stewart, T. P. *Nature* **290**, 427-428 (1980).
5. Rand, R. P., Reese, T. S. & Miller, R. G. *Nature* **293**, 237-238 (1981).
6. Tokutomi, S., Lew, R. & Ohnishi, S. *Biochim. biophys. Acta* **643**, 276-282 (1981).
7. Ohnishi, S. & Tokutomi, S. in *Biological Magnetic Resonance* Vol. 3 (eds Berliner, L. & Reuben, J.) 121-153 (Plenum, New York, 1980).
8. Sackmann, E. *Ber. Bunsenges. phys. Chem.* **82**, 891-909 (1978).
9. Berclaz, T. & McConnell, H. M. *Biochemistry* **20**, 6635-6640 (1981).
10. Servuss, R. M., Harbich, W. & Helfrich, W. *Biochim. biophys. Acta* **436**, 900-903 (1976).
11. Helfrich, W. *Z. Naturf.* **28c**, 693-703 (1973).

Effects of Duchenne muscular dystrophy on muscle protein synthesis

M. J. Rennie*, R. H. T. Edwards*, D. J. Millward†, S. L. Wolman‡, D. Halliday‡ & D. E. Matthews§

* Department of Human Metabolism, School of Medicine, University College London, Rayne Institute, University Street, London WC1E 6JJ, UK

† Department of Human Nutrition, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

‡ Division of Clinical Sciences, Clinical Research Centre, Northwick Park Hospital, Harrow HA1 3UJ, UK

§ Metabolism Division, Washington University School of Medicine, St Louis, Missouri 63110, USA

The primary defect in Duchenne muscular dystrophy (DMD) is unknown but muscle growth failure and wasting of skeletal muscle must be the result of an imbalance between protein synthesis and degradation, that is, the rate of muscle protein synthesis is insufficient to replace degraded proteins and expand the protein mass. In patients with DMD, excretion of 3-methylhistidine is dramatically increased relative to creatinine and this is commonly accepted as evidence of increased muscle myofibrillar degradation¹⁻³. As rapid wasting of muscle has not been observed to occur in the disease, increased protein degradation ought to be accompanied by increased protein synthesis and this has been observed or implied from measurements in animal models of the disease⁴⁻⁶. However, so far no direct measurements of muscle protein synthesis or degradation have been made in patients. We report here the first direct measurements *in vivo* of rates of skeletal muscle protein synthesis in patients suffering from DMD. The results, obtained using stable isotope tracers, clearly show that a marked reduction in muscle protein synthesis is the primary cause of the growth failure and that muscle protein degradation cannot be elevated.

We studied seven normal fed adult men aged 22-65 yr and five boys with Duchenne muscular dystrophy, aged 12-18 yr. Three of the patients were wheelchair bound and the two youngest were on the verge of becoming so. Investigations of these subjects using radioactive isotopes would not have been ethically feasible. We infused $[1-^{13}\text{C}]$ leucine intravenously after a priming dose⁷, and measured the incorporation of labelled leucine into skeletal muscle proteins in samples obtained by needle biopsy⁸ (Fig. 1). This approach also allows the measurements of whole-body leucine flux and its components, leucine oxidation and whole-body protein synthesis⁹ (Table 1).

The rate of muscle protein synthesis in normal adults was determined from the increase in isotopic enrichment between two biopsies; as synthesis rates calculated from the enrichment observed in a single biopsy taken after 7 h were not significantly different from this estimate, we calculated synthesis rates in patients from the enrichment in a single biopsy sample taken at 7.5 h. The precursor labelling was assumed to be that of the keto acid of leucine, α -ketoisocaproate (α -KIC), in plasma.

Most plasma KIC originates from skeletal muscle¹⁰ so that its enrichment should closely reflect that of intracellular leucine in muscle. This was confirmed in four normal adults and three patients by direct measurements of the enrichment of free leucine in the muscle biopsy intracellular water (unpublished results).

In normal adults (Table 1) the rate of muscle protein synthesis was 0.198% h⁻¹, twice the value determined previously¹¹. However, in those studies using ¹⁵N-lysine, the precursor labelling was taken as that of plasma lysine, which we believe gave erroneous results. Studies using infusions of lysine and leucine in pigs have shown that the ratio of intramuscular free lysine labelling relative to plasma is less than 50% of that for leucine¹², which explains the discrepancy and supports our current measurements. From creatinine excretion as a proportion of muscle mass¹³ and from our measurements of muscle protein content (17%) and leucine content of protein (8%), we estimate that in a man weighing 70 kg, 126 g of protein are synthesized per 12 h in the fed state, that is 53% of the whole-body synthesis rate. This is a higher proportion than that observed in smaller animals¹⁴.

In the patients with DMD, whole-body protein synthesis was depressed as was that in muscle (Table 1), the former being only half that observed in normal adults, and as all dietary leucine was oxidized, no net protein deposition could occur. The rate of muscle protein synthesis measured during feeding

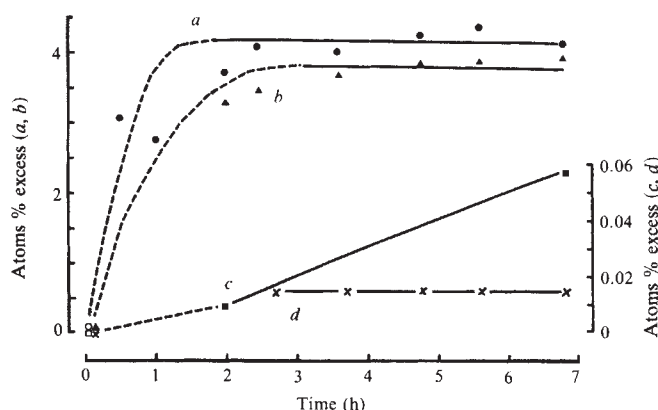


Fig. 1 Enrichment with ¹³C of blood leucine (●) and of α-KIC (▲) (transamination product of leucine), leucine incorporated into quadriceps muscle protein (■) (obtained by needle biopsy⁸, with local anaesthesia) and of expired CO₂ (x) during the primed (0.8 mg per kg) infusion (1 mg per kg per h) of [1-¹³C]leucine (KOR Inc., Massachusetts) in a normal healthy adult man (80 kg, 42 yr old). A similar pattern of changes was observed in all subjects. The subject was fed 0.1 g protein and 14 kJ per kg at hourly intervals. Blood was sampled via an indwelling cannula and the enrichment with ¹³C of leucine and α-KIC determined by gas chromatography-mass spectrometry^{7,30}. Expired CO₂ was collected in 5-l Douglas bags attached to a three-way valve and mouthpiece and a portion of the sample immediately transferred to an evacuated glass vial. The enrichment of CO₂ with ¹³C was measured in a VG micromass 602D isotope ratio mass spectrometer after drying and trapping of solid CO₂ in a cold-finger cooled with liquid N₂ (ref. 31). Total CO₂ production was measured from the IR absorbance of expired air sampled continuously from a metabolic tent, or collected over 5 min in 100-l Douglas bags. Alkali-soluble muscle protein from fat-free, dried biopsy samples was hydrolysed and leucine was separated preparatively on a Locarte amino acid analyser. Leucine-CO₂ was liberated by ninhydrin at pH 2, dried and its enrichment measured in the isotope ratio mass spectrometer. Muscle protein synthesis, k_s , may be obtained by comparing the change in ¹³C enrichment of leucine incorporated into protein in two muscle biopsy samples taken some hours apart, but at times (t_1 , t_2) when the blood leucine or α-KIC pools are enriched to plateau values. Synthesis rate, $k_s = \Delta E_{\text{muscle}} / \sum_{i=1}^{t_2} E_{\text{blood}} \Delta t \times 100 \text{ h}^{-1}$ where ΔE_{muscle} is the change in enrichment of muscle and $\sum E_{\text{blood}}$ is the integral of blood over the time in h. Note in the figure that with an assumed initial zero (that is, background) enrichment of muscle-protein leucine (as measured in muscle from surgical patients, in mixed blood proteins and in casein), incorporation of ¹³C into muscle protein is essentially linear with time. There is thus little error in calculating synthesis on the basis of either the change in enrichment between the biopsies compared with the plateau of blood leucine or α-KIC enrichment, or the incorporation of ¹³C-leucine into muscle obtained by a single biopsy at the end of the whole period of infusion, compared with the integrated weighted-mean blood leucine (or α-KIC).

Table 1 Skeletal muscle and whole-body protein synthesis in normal adults and boys with Duchenne muscular dystrophy

	Normal adults (n = 7)	Boys with DMD (n = 5)
Whole-body protein turnover		
Body weight (kg)	80.8 ± 19.2	54 ± 6.5*
Leucine flux $\dot{Q}$ (μmol per kg per h)	194 ± 20	140 ± 39
Leucine intake $\dot{I}$ (μmol per kg per h)	55 ± 4	64 ± 10
Leucine oxidation $\dot{O}$ (μmol per kg per h)	40 ± 8	63 ± 30*
Protein synthesis $\dot{S}$ (μmol leucine per kg per h)	154 ± 30	77 ± 22*
(g protein per 12 h)	212 ± 41	81.7 ± 23*
Muscle protein synthesis		
Synthesis rate (% h ⁻¹)	0.198 ± 0.055	0.066 ± 0.028*
Muscle mass (kg)	31.1 ± 7.7	6.4 ± 3.0*
Muscle protein synthesis (g per 12 h)	113 ± 31	17.5 ± 7.42*
Urinary 3-methylhistidine/creatinine (molar ratio × 10 ³)	19.0 ± 2.4	55.2 ± 5.4*

Data are mean ± s.d. for volunteers, consisting of seven healthy adult males (aged 22–65 yr) and five children with Duchenne muscular dystrophy, aged 12–18 yr. Subjects were infused intravenously with carboxyl-labelled ¹³C-leucine (enriched to 90 atoms % excess) for 7–8 h of the study while being fed a liquid milk-based food at hourly intervals (~0.1 g, protein, 14 kJ per kg body wt). Leucine flux and oxidation were calculated using equations given in ref. 7, except that the enrichment of blood α-ketoisocaproate rather than leucine was used as this is more likely to represent the labelling of the metabolic pool from which leucine is removed by protein synthesis and oxidation. The component of the flux not directly measured, protein synthesis ($\dot{S}$), was obtained by manipulation of the data for flux ($\dot{Q}$) and oxidation rate ($\dot{O}$), according to a two-pool stochastic model whereby in the steady state $\dot{Q} = \dot{S} + \dot{O}$ (that is, the sum of processes removing leucine)⁸. (See refs 7, 30, 31 for details of mass spectrometric analysis.) Muscle synthesis rate was measured by ¹³C incorporation into leucine of mixed muscle protein of quadriceps (see Fig. 1) assuming that the precursor pool for protein synthesis was labelled similarly to α-KIC in blood. Urinary 3-methylhistidine and creatinine were measured in most normal subjects over 24 h on various occasions but were also measured during infusion of labelled leucine. In the case of the patients with muscle disease, values for urinary 3-methylhistidine and creatinine are the mean of estimates obtained over a number of days. There were no significant differences in the mean values of such estimates compared with those obtained during the period of infusion. Muscle mass was estimated from creatinine excretion (8.84 mmol of creatinine = 20 kg of muscle)¹³.

* Significant difference ($P < 0.01$; Student's one-tailed t -test) from levels in normal fed men.

was only 34% of that observed in the normal adults (that is, 0.066% h⁻¹), which is less than that observed during fasting in normal adults (unpublished results). This means that muscle protein synthesis contributes only 17.5 g protein per 12 h, 8.6% of whole-body protein synthesis. This depression of muscle protein synthesis is all the more marked when it is considered that in normal adolescent children who are still growing, the rates of whole-body and muscle protein synthesis are likely to be much higher than in adults.

The finding of a depressed rate of muscle protein synthesis associated with DMD was unexpected, particularly in view of the high 3-methylhistidine/creatinine ratio observed in the present patients and reported previously in other patients with this disease^{1–3}. 3-Methylhistidine occurs in actin and some species of myosin heavy chain, and as it is not metabolized in man its excretion should be a quantitative estimate of the rate of actin and myosin heavy chain degradation¹⁵. Creatinine excretion may be taken as an index of skeletal muscle mass¹³ (even in cases of muscle disease¹⁶); thus the threefold elevated 3-methylhistidine/creatinine ratio observed in boys with DMD has been taken to indicate a proportionally elevated rate of myofibrillar degradation^{1–3}. In addition, as these patients were not actually losing muscle protein at a significant rate (on the basis of our repeated computerized tomography and creatinine excretion measurements), the rate of protein synthesis should therefore be increased compared with normal values. Clearly this conclusion does not agree with the rates of muscle protein synthesis measured here, and have no reason to believe that our direct measurements of the rate of muscle protein synthesis are substantially wrong. For this to be the case, enrichment of the true precursor leucine pool would have to be very much lower than the α-KIC enrichment in the patients compared

with the normal adult subjects, which is unlikely. On the other hand, we have already reported findings that raise doubts about the validity of the assumption that 3-methylhistidine excretion rates are a true reflection of skeletal muscle degradation rates¹⁷⁻¹⁹. We have shown that in the rat, substantial amounts of 3-methylhistidine may derive from the degradation of actin in non-muscle tissues. Although the amounts of actin in tissues other than skeletal muscle is small, it is a ubiquitous protein and if its turnover rate in non-muscle tissues is high this could result in a substantial non-muscle contribution to the urine.

In fact the proportion of non-muscle sources of excreted 3-methylhistidine is likely to be much higher in the dystrophic patients than in normal subjects. We find, on the basis of creatinine excretion, body weight and total body potassium measurements, that although skeletal muscle mass is markedly reduced in these patients (to only 12% of body weight), non-skeletal muscle mass (normally 15% of body weight) is only reduced to 74% of normal. If creatinine excretion is depressed in proportion to the reduced muscle mass, non-muscle sources of 3-methylhistidine would inflate the 3-methylhistidine/creatinine ratio, as has been observed in muscle wasting due to amyopathic lateral sclerosis²⁰. Given the 3-methylhistidine content of skin ($0.6 \mu\text{mol per g protein}$) and gastrointestinal muscle ($1.8 \mu\text{mol per g protein}$)²¹ and assuming that these two tissues comprise a normal proportion (47 and 13% respectively)^{22,23} of non-skeletal muscle tissues in patients with muscular dystrophy, then these tissues need only turn over at three to four times the rate of skeletal muscle to account for the excreted 3-methylhistidine. Studies of patients undergoing elective surgery have provided evidence that in man, skin and gut muscle do indeed turn over at about two and five times, respectively, the rate of skeletal muscle²⁴.

An elevated rate of protein degradation in diseased muscle has also been observed in animal models of the disease⁴⁻⁶. However, there is no animal model which reproduces exactly the features of the human disease. Indeed, in most of the animal models there is a transient muscle hypertrophy early on in the disease, and protein turnover measurements are most often made in animals at this stage. As we have pointed out elsewhere²⁵, increased protein turnover in muscle is often a feature of rapid growth and regeneration, which may account for the elevated turnover in these animal models²⁶. There is no evidence to suggest that this occurs in adolescent boys with muscular dystrophy such as those studied here, although it may occur in very much younger patients with the disease.

A reduced rate of protein synthesis rather than an increased rate of protein degradation as the primary cause of growth failure in DMD has important implications for present disease therapies. The notion that increased degradation was the main factor led to suggestions (and testing in animal models) that proteinase inhibitors might prove effective²⁷⁻²⁹. However, suppression of muscle protein degradation falsely assumed to be elevated in patients would be of little value since protein synthesis is markedly depressed. The use of possible therapeutic agents to increase muscle protein synthesis would be much more rational. The present results demonstrate that it would be possible to monitor such effects over the short term.

This work was supported by grants from the Muscular Dystrophy Group of Great Britain, Muscular Dystrophy Association of America, the British MC, Wellcome Trust, Action Research for the Crippled Child, Abbott Laboratories, and NIH grants AM-25994 and RR-00954. S.L.W. was a Canadian MRC Fellow. We thank the dietitians and nursing staff of the Hospital for Tropical Disease and University College Hospital for their assistance and the staff of the London Broadcasting Company's A.M. programme who helped us locate lost samples. We also thank M. A. Read, W. W. Read, M. Nathan, P. C. Bates and A. Mellone for technical assistance.

Received 14 October 1981; accepted 26 January 1982.

- McKernan, R. O., Halliday, D. & Purkiss, P. J. *Neurol. Neurosurg. Psychiatr.* **40**, 979-981 (1977).
- Ballard, F. J., Tomas, F. M. & Stern, L. M. *Clin. Sci.* **56**, 347-352 (1979).
- Warnes, D. M., Tomas, F. M. & Ballard, F. J. *Muscle Nerve* **4**, 2-11 (1981).

- Garber, A. J., Schwartz, R. J., Seidel, C. L., Silvers, A. & Entman, M. L. *J. biol. Chem.* **255**, 8315-8324 (1980).
- Li, J. B. *Am. J. Physiol.* **239**, 401-406 (1981).
- Hillgartner, F. B., Williams, A. S., Flanders, J. A., Morin, D. & Hansen, R. J. *Biochem. J.* **196**, 591-601 (1980).
- Matthews, D. E. *et al. Am. J. Physiol.* **238**, E473-E481 (1980).
- Edwards, R., Young, A. & Wiles, M. *New Engl. J. Med.* **302**, 261-271 (1980).
- Waterlow, J. C., Garlick, P. J. & Millward, D. J. *Protein Turnover in Mammalian Tissues and in the Whole Body* (Elsevier, Amsterdam, 1978).
- Adibi, S. A. *Metabolism* **25**, 1287-1302 (1976).
- Halliday, D. & McKernan, R. O. *Clin. Sci. molec. Med.* **49**, 581-590 (1975).
- Simon, O., Munchmeyer, R., Bergner, H., Zebrowska, T. & Buraczewska, L. *Br. J. Nutr.* **40**, 243-252 (1978).
- Graystone, J. E. in *Human Growth* (ed. Cheek, D. B.) 182-197 (Lea & Febiger, Philadelphia, 1968).
- Millward, D. J. *Reprod. Nutr. Dév.* **21**, 265-277 (1981).
- Young, V. R. & Munro, H. N. *Fedn Proc.* **37**, 2291-2297 (1979).
- Murray, C. E., Warnes, D. M., Ballard, F. J. & Tomas, F. M. *Clin. Sci.* **61**, 737-741 (1981).
- Millward, D. J. *et al. Biochem. J.* **190**, 225-228 (1980).
- Bates, P. C. & Millward, D. J. *Proc. Nutr. Soc.* **40**, 89A (1981).
- Millward, D. J., Bates, P. C., Broadbent, P. & Rennie, M. J. *Proc. 3rd Congr. European Society on Parenteral and Enteral Nutrition* (ed. Wesdorp, E. C.) (Churchill-Livingstone, Edinburgh, in the press).
- Afting, E. G., Bernhardt, W., Janzen, R. W. C. & Röthig, H. J. *Biochem. J.* **220**, 449-452 (1981).
- Elia, M., Carter, A. & Smith R. *Br. J. Nutr.* **42**, 567-570 (1979).
- Jackson, C. M. *Am. J. Anat.* **9**, 119-124 (1909).
- Forbes, B. M., Cooper, A. R. & Mitchell, H. H. *J. biol. Chem.* **203**, 359-367 (1953).
- Read, M. A. *et al. Proc. 3rd Congr. European Society on Parenteral and Enteral Nutrition* (ed. Wesdorp, R. C.) (Churchill-Livingstone, Edinburgh, in the press).
- Millward, D. J. in *Degradative Processes in Skeletal and Cardiac Muscle* (ed. Wiidenthal K.) 161-200 (Elsevier, Amsterdam, 1980).
- Rennie, M. J., Edwards, R. H. T. & Millward, D. H. *Muscle Nerve* (in the press).
- McGowan, E. P., Shafiq, S. A. & Stracher, A. *Expl Neurol.* **50**, 649-652 (1976).
- Stracher, A., McGowan, E. B. & Shafiq, S. A. *Science* **200**, 500-501 (1978).
- Libby, P., Ingwall, J. G. & Goldberg, A. L. *Am J. Physiol.* **237**, E35-E39 (1979).
- Matthews, D. E., Ben-Galim, E. & Bier, D. M. *Analyt. Chem.* **51**, 80-88 (1979).
- Halliday, D. & Read, W. W. C. *Proc. Nutr. Soc.* **40**, 321-334 (1981).

High-affinity ouabain binding site and low-dose positive inotropic effect in rat myocardium

Robert J. Adams, Arnold Schwartz, Gunther Grupp, Ingrid Grupp, Shin-Woong Lee & Earl T. Wallick

Departments of Pharmacology and Cell Biophysics, Physiology and Medicine, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267, USA

Trevor Powell, Victor W. Twist & Premjith Gathiram

Department of Physics as Applied to Medicine, Middlesex Hospital Medical School, London WIP 7PN, UK

The mechanism of the positive inotropic effect of digitalis glycosides remains unclear. One theory suggests a causal relationship between the binding to and consequent inhibition of $(\text{Na}^+ + \text{K}^+)$ ATPase (the sodium pump) by digitalis and an increased myocardial contractile force¹⁻⁶. By this mechanism, the increased force of contraction would occur secondary to an elevation of intracellular sodium concentration which then causes an increased intracellular concentration of calcium via a sodium-calcium exchange mechanism⁴⁻⁶. Another theory proposes that the binding of digitalis to $(\text{Na}^+ + \text{K}^+)$ ATPase causes an increase in a sarcolemmal calcium pool⁶⁻⁹, suggesting that the inotropic effect could be due to a causal relationship between the formation of the digitalis- $(\text{Na}^+ + \text{K}^+)$ ATPase receptor complex and increased myocardial calcium availability and utilization, exclusive of an inhibition of the sodium pump⁷⁻⁹. We now report that two distinct positive inotropic sites for ouabain exist in rat ventricular strips. The higher-affinity response ($\text{ED}_{50} = 0.5 \mu\text{M}$) correlates with an apparent high-affinity site which can be detected by ³H-ouabain binding to intact rat ventricular myocytes. These higher-affinity sites do not correlate with concentrations (IC_{50}) of ouabain necessary to inhibit $(\text{Na}^+ + \text{K}^+)$ ATPase activity of sarcolemma preparations prepared from rat ventricles, suggesting that in the rat ventricle the high-affinity site for the inotropic effect of ouabain may not be related to inhibition of $(\text{Na}^+ + \text{K}^+)$ ATPase. The low-affinity site is, however, related to inhibition of this enzyme.

Figure 1 shows a typical cumulative dose-response curve of ouabain on isometric contractile force in rat right ventricular strips. In a series of identical experiments, a 'low concentration

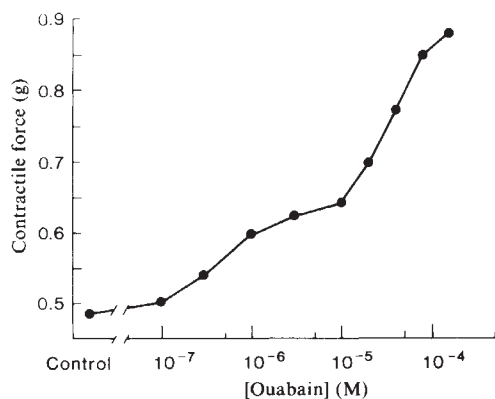


Fig. 1 Typical biphasic effect of ouabain (in mol l^{-1}) on isometric contractile force (in g) in rat ventricular myocardial strips. Right ventricular strips were prepared from adult rats as previously described¹¹ and vertically mounted in a water-jacketed muscle bath containing (in mM) 137 NaCl, 5.4 KCl, 1.8 CaCl_2 , 1.05 MgSO_4 , 0.4 KH_2PO_4 , 12 NaHCO_3 and 10 glucose. This Tyrode solution was aerated with 95% O_2 -5% CO_2 and maintained at a pH of 7.0 at 35 °C. The strips were allowed to equilibrate for 60 min while electrically paced at 1 Hz (via punctate platinum electrodes) with a stimulus duration of 2 ms and voltage at 20% above threshold (6–8 V). During this equilibration period, the preload was constantly adjusted to 1 g. A single cumulative concentration-response curve for ouabain was obtained for each ventricular strip. For the experiment shown, the apparent ED_{50} s were estimated to be 0.5 and 19 μM .

effect' occurred between 0.1 and 10 μM ouabain, with an increase in contractile force of $34 \pm 5\%$ ($n = 25$) above control, while a 'high concentration effect', with a maximal increase of $84 \pm 9\%$ ($n = 25$) above control contractile force, occurred with 80–160 μM ouabain. Although the increase in contractility was undoubtedly produced by a complex series of events, the obvious biphasic nature of the ouabain dose-response curve (Fig. 1) is consistent with the existence of two separate glycoside receptor sites. We have therefore estimated two ED_{50} values from the experimental data. The high-affinity site had an ED_{50} of $0.53 \pm 0.08 \mu\text{M}$ (assuming that the inflection point in the curve at 10^{-5} M represents the maximum effect of this site) and the low-affinity site had an ED_{50} of $19 \pm 2 \mu\text{M}$. Neither β -adrenoreceptor blockade (20 μM sotalol) nor pretreatment of the rats with reserpine had any effect on the biphasic nature of the curve, the maximal increases in contractile force or the apparent ED_{50} values for the low or high concentration effects, but the low-dose effect could be removed by prolonged washing. Erdmann *et al.*¹⁰ recently reported that rat ventricular strips had a single ouabain inotropic receptor with an ED_{50} of 0.3 μM . This value is lower than found by us¹¹ and may be due to the fact that Erdmann *et al.*¹⁰ were unable fully to elicit the high concentration effect, because in their preparations toxicity occurred at ouabain concentrations $>30 \mu\text{M}$.

In the same experimental conditions used to study rat ventricular strips, we obtained only a monophasic concentration-response curve for rat left atria ($\text{ED}_{50} = 45 \pm 9 \mu\text{M}$), which is consistent with earlier reports of the positive inotropic effects of digitalis glycosides in rat myocardium¹¹. Similarly, only monophasic response curves were obtained for cat, rabbit and guinea pig ventricular papillary muscles or left atria. Thus, two inotropic receptor sites for ouabain seem to exist in rat ventricle that are perhaps less easily detectable in rat atria or myocardium of the other species.

Although most studies of ^3H -ouabain binding to various membrane preparations and purified $(\text{Na}^+ + \text{K}^+)\text{ATPases}$ reveal a single class of ouabain binding site¹², several workers have recently obtained results consistent with the existence of two populations of such sites^{11,13–17}. In an effort to detect two receptor sites of different affinities, we elected to examine ^3H -ouabain binding to isolated myocytes^{18,19} instead of to membrane preparations, for two main reasons. First, it is likely that cell membrane preparations from ventricular myocardium contain $(\text{Na}^+ + \text{K}^+)\text{ATPase(s)}$ from more than one cell type, including fibroblasts and postganglionic sympathetic nerves, which probably have different affinities for ouabain, particularly in

the rat¹⁷. Second, the preparative procedure for cardiac sarcolemma can potentially alter the configuration of all membranes and thus artificially create more than one apparent population of receptors.

The method of Akera and Cheng²⁰ was used to examine ^3H -ouabain binding to isolated rat ventricular myocytes. A K_D value can be calculated from data obtained in a displacement assay (Fig. 2) by subtracting the concentration of the labelled ouabain (a) from the concentration of unlabelled ouabain ($C_{0.5}$) which causes a 50% displacement of the labelled ligand. Using concentrations of labelled ouabain of 0.1 μM ($n = 3$) and 0.2 μM ($n = 3$), we obtained from probit plots $C_{0.5}$ values of $0.225 \pm 0.066 \mu\text{M}$ and $0.322 \pm 0.028 \mu\text{M}$, respectively, to give an average $K_D \pm \text{s.d.}$ of $0.124 \pm 0.090 \mu\text{M}$. Akera and Cheng²⁰ emphasize that this calculation of K_D applies only to a single class of site and that the binding reaction must be at equilibrium. Although equilibrium is difficult to obtain in ouabain-sensitive species, which have slow dissociation rates^{5,21,22}, equilibrium binding of ^3H -ouabain to rat cardiac $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ can be achieved rapidly due to a high dissociation rate^{5,21}. We found 30 min sufficient for equilibrium binding in our studies (data not shown). Because of large variability in the data obtained at the higher concentrations of unlabelled ouabain (where labelled ligand binding is low) we could not determine whether or not there was a second, lower-affinity site for ouabain binding. Such a low-affinity site clearly exists, as the IC_{50} for inhibition of our rat cardiac sarcolemma preparation is 50 μM , similar to previously reported values for rat heart $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ ^{11,21–25}. This low-affinity site for enzyme inhibition agrees well with that for positive inotropy (Fig. 1). Thus, our results are not consistent with the conclusion of Erdmann *et al.*¹⁰ that there is a complete dissociation of the positive inotropic effect of ouabain from the low-affinity site.

Studies reported by Godfraind and Ghysel-Burton^{26–28} on ouabain uptake, tissue electrolyte content and contractile force in guinea pig atria, and experiments reviewed by Noble²⁹ on

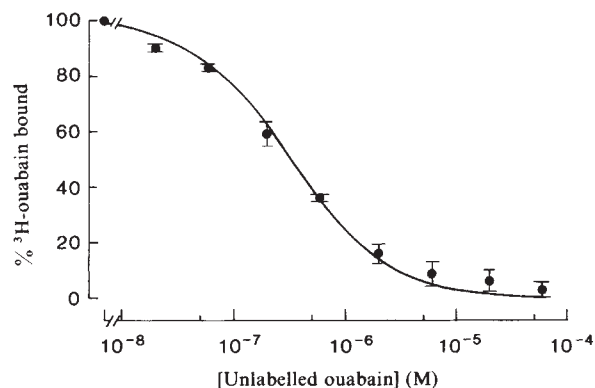


Fig. 2 ^3H -ouabain binding to intact ventricular myocytes, isolated from adult rat myocardium as previously described^{18,19}. The curve shows the effect of increasing amounts of unlabelled ouabain on the binding of 0.2 μM ^3H -ouabain (100% binding) to myocytes in the presence of Krebs-Ringer bicarbonate buffer containing (in mM) 118.5 NaCl, 2.6 KCl, 0.5 CaCl_2 , 1.18 MgSO_4 , 1.18 KH_2PO_4 , 14.5 NaHCO_3 , 11.1 glucose and 2.0% (w/v) bovine serum albumin (Miles fraction V); the final incubation volume was 2 ml. To each assay tube (5 ml Teflon beakers) 200 μl of ^3H -ouabain (1 μM , 17 Ci mmol^{-1} in aqueous solution) were added with appropriate amounts of unlabelled ouabain (stock solutions prepared in Krebs-Ringer bicarbonate buffer). After preincubation for 15 min in a shaking water bath (37 °C) aerated with 95% O_2 and 5% CO_2 in a closed hood, the binding reaction was initiated by the addition of 0.5 ml of cells ($\sim 4 \times 10^4$ cells). The reaction was terminated after 30 min by transferring the incubation beakers to an ice bath, followed by rapid filtration on Whatman GF/C glass fibre filters. The beakers were rinsed with 10 ml cold KCl (100 mM) and this rinse was used to wash the filters (filtration and wash time was ~ 5 s). The filters were placed in vials and incubated for at least 10 h in scintillation fluid before determination of radioactivity. External standards were prepared using ^3H -ouabain and nonspecific binding was defined as the bound radioactivity not displaced by $5.5 \times 10^{-3} \text{ M}$ ouabain, which was $\sim 30\%$ of the total binding in the absence of any added unlabelled ouabain. The curve shown is predicted from the equation, $\% ^3\text{H}\text{-ouabain bound} = 100 / (1 + C/C_{0.5})$ where C = concentration of unlabelled ouabain and $C_{0.5} = K_D + a$, with $a = 0.2 \mu\text{M}$ and $K_D = 0.125 \mu\text{M}$. Each point represents the mean ($\pm \text{s.e.m.}$) of three separate assays with a concentration of 0.2 μM ^3H -ouabain.

potassium-sensitive reversal potentials, current-voltage relations and twitch tension in sheep Purkinje fibres, have suggested the presence of two ouabain binding sites of high and low affinity, related to Na^+, K^+ -pump stimulation and inhibition, respectively. Low doses of ouabain ($\sim 10^{-8}$ M) can often produce a small but reversible negative inotropic effect, which may be transient^{29,30}. Our results, obviously consistent with the presence of two ouabain binding sites, show a reproducible positive inotropic effect of ouabain at levels corresponding to binding of the glycoside to a high-affinity site. This positive inotropy cannot be causally dependent on an overall decrease in sodium gradient if the direct or indirect effects of low doses of ouabain are to produce a net stimulation of the Na^+, K^+ -pump²⁹.

The striking discrepancy between the apparent IC_{50} value for ouabain inhibition of rat heart ($\text{Na}^+ + \text{K}^+$)ATPase and the K_D of the higher-affinity ouabain binding site reported here, suggests that there may be no relationship between Na^+, K^+ -pump inhibition and the low concentration inotropic effect in rat ventricle. However, the high-affinity ouabain binding site may represent only a small fraction of the total number of ($\text{Na}^+ + \text{K}^+$)ATPase molecules in rat ventricle, the majority being of the lower-affinity type. In this case, inhibition at the higher-affinity site by low concentrations of ouabain may produce a limited increase in internal sodium (and secondarily of calcium), because the pump still has sufficient 'gain' to keep internal sodium relatively low. At higher glycoside concentrations, binding to the low-affinity site reduces pump gain and internal sodium will rise rapidly. The exact shape of the force-binding curve will depend on the relationship between internal sodium and force, which is not simple³¹, but a biphasic response would not be surprising. This hypothesis is difficult to test experimentally, as we cannot detect a change in rat cardiac ($\text{Na}^+ + \text{K}^+$)ATPase activity at ouabain concentrations below 1 μM . Alternatively, the higher-affinity site may exist as a catalytically inactive conformation or isozyme of ($\text{Na}^+ + \text{K}^+$)ATPase. In this case, high-affinity binding of ouabain *in vivo* would not result in pump inhibition, but perhaps induce a conformational change in the enzyme (and enzyme-associated membrane lipids) which could increase calcium binding to sarcolemma^{7,9,15}.

Thus, ouabain binding to the higher-affinity ($\text{Na}^+ + \text{K}^+$)ATPase in rat ventricular sarcolemma could result in increased calcium mobilization and availability to the myofibrils, producing the low-concentration positive inotropic effect in isolated rat ventricular muscle. These effects could occur independently of Na^+, K^+ -pump inhibition. With increasing concentrations of ouabain, binding to the lower-affinity site would occur, resulting in Na^+, K^+ -pump inhibition, enhanced $\text{Na}^+ - \text{Ca}^{2+}$ exchange and a more pronounced positive inotropic effect. While there appears to be a clear distinction between the two sites as well as between the two inotropic phases in rat ventricular myocardium, these differences in myocardium of ouabain-sensitive species may be small and thus difficult to detect if, in fact, two populations of ouabain sites actually exist.

This work was supported by USPHS grants PO1 HL 22619 and F32 HL 05802, the American Heart Association, Southwestern Ohio Chapter, the British Heart Foundation and the MRC.

Received 17 August 1981; accepted 21 January 1982.

1. Repke, K. H. R. in *New Aspects of Cardiac Glycosides* (ed. Wilbrandt, W.) 47-73 (Pergamon, London, 1963).
2. Besch, H. R. Jr, Allen, J. C., Glick, G. & Schwartz, A. *J. Pharmac. exp. Ther.* **171**, 1-12 (1970).
3. Akera, T., Larsen, F. S. & Brody, T. *J. Pharmac. exp. Ther.* **173**, 145-151 (1970).
4. Schwartz, A., Lindenmayer, G. E. & Allen, J. C. *Pharmac. Rev.* **27**, 3-134 (1975).
5. Akera, T. & Brody, T. M. *Pharmac. Rev.* **29**, 187-220 (1978).
6. Schwartz, A. & Adams, R. J. *Circulation Res.* **46**, 1154-1160 (1980).
7. Gervais, A., Lane, L. K., Anner, B. M., Lindenmayer, G. E. & Schwartz, A. *Circulation Res.* **40**, 8-14 (1977).
8. Busse, F., Lullmann, H. & Peters, T. *J. cardiovascular pharmacol.* **1**, 687-698 (1979).
9. Lullmann, H. & Peters, T. *Prog. Pharmac.* **2**, 1-57 (1979).
10. Erdmann, E., Philipp, G. & Scholz, H. *Biochem. Pharmac.* **29**, 3219-3229 (1980).
11. Grupp, G., Grupp, I., Johnson, C. L., Wallick, E. T. & Schwartz, A. *Biochem. biophys. Res. Commun.* **88**, 440-447 (1979).
12. Wallick, E. T., Lane, L. K. & Schwartz, A. *Rev. Physiol.* **41**, 397-411 (1979).

13. Fricke, U. & Klaus, W. *Br. J. Pharmac.* **61**, 23-428 (1977).
14. Heller, M. & Beck, S. *Biochim. biophys. Acta* **514**, 332-347 (1978).
15. Wellsmith, N. V. & Lindenmayer, G. E. *Circulation Res.* **47**, 710-720 (1980).
16. Onji, T. & Liu, M.-S. *Archs Biochem. Biophys.* **207**, 148-156 (1981).
17. Swadner, K. J. *J. biol. Chem.* **254**, 6060-6067 (1979).
18. Powell, T. & Twist, V. W. *Biochem. biophys. Res. Commun.* **72**, 327-333 (1976).
19. Powell, T., Terrar, D. A. & Twist, V. W. *J. Physiol., Lond.* **302**, 131-153 (1980).
20. Akera, T. & Chang, V.-J. K. *Biochim. biophys. Acta* **470**, 412-423 (1977).
21. Allen, J. C. & Schwartz, A. *Pharmac. exp. Ther.* **168**, 42-46 (1969).
22. Wallick, E. T., Pitts, B. J. R., Lane, L. K. & Schwartz, A. *Archs Biochem. Biophys.* **202**, 442-449 (1980).
23. Sharma, V. K. & Banerjee, S. P. *Molec. Pharmac.* **14**, 122-129 (1978).
24. Repke, K., Est, M. & Portius, H. J. *Biochem. Pharmac.* **14**, 1785-1802 (1965).
25. Akera, T., Larsen, F. S. & Brody, T. M. *J. Pharmac. exp. Ther.* **170**, 17-26 (1969).
26. Godfraind, T. & Ghysel-Burton, J. *Nature* **265**, 165-166 (1977).
27. Ghysel-Burton, J. & Godfraind, T. *Br. J. Pharmac.* **66**, 175-184 (1979).
28. Godfraind, T. & Ghysel-Burton, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3067-3069 (1980).
29. Noble, D. *Cardiovascular Res.* **14**, 495-514 (1980).
30. Grupp, G., Grupp, I., Schwartz, A., Ghysel-Burton, J. & Godfraind, T. *J. Pharmac. exp. Ther.* (in the press).
31. Eisner, D. A., Lederer, W. J. & Vaughan-Jones, R. D. *J. Physiol., Lond.* **317**, 163-187 (1981).

Nanomolar concentrations of extracellular ATP activate membrane Ca^{2+} channels in snail neurones

A. Yatani, Y. Tsuda, N. Akaike & A. M. Brown*

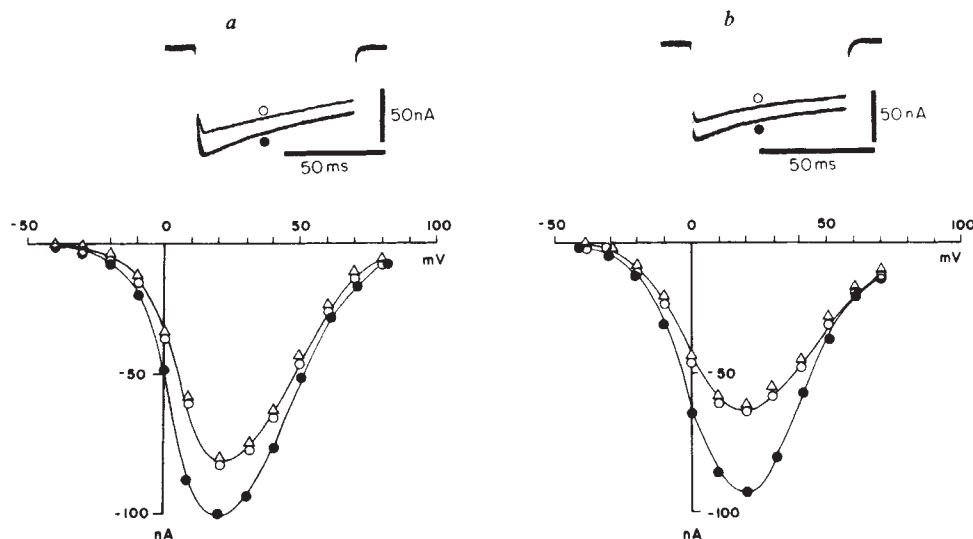
Department of Physiology and Biophysics, University of Texas Medical Branch, Galveston, Texas 77550, USA

Intracellular ATP is known to affect the electrical properties of membranes indirectly by acting intracellularly as a substrate for ATPases such as ($\text{Na}^+ + \text{K}^+$)ATPase, Ca^{2+} -ATPase or adenylyl cyclase. More recently, extracellular ATP has been shown to act on effector cells innervated by autonomic nerves and a purinergic receptor hypothesis has been proposed for this effect^{1,2}. Extracellular ATP also acts on the slow inward current, I_{ss} , of cardiac muscle membranes^{3,4} where the effect may be coupled to a change in intracellular metabolism. These observations raise the possibility that extracellular ATP (ATP_o) might alter the Ca current of excitable membranes. The Ca current is of major interest because it acts as the trigger for many cellular functions such as secretion, contraction, motility and luminescence. To test our idea, we examined the effects of ATP_o on a well studied Ca current, that of neuronal membranes. As we report here, ATP_o in nanomolar to micromolar concentrations produced substantial increases of I_{Ca} , and the non-hydrolysable congener of ATP, adenylylimidodiphosphate (AMP-PNP) had an even greater effect. By contrast, intracellular perfusion of ATP in nanomolar quantities had no effects on I_{Ca} , but with 10^{-5} - 10^{-3} MATP small increases were observed. Intracellular perfusion with AMP-PNP at these doses had no effect on I_{Ca} . We conclude that ATP_o enhances membrane Ca channel activity by an as yet undetermined mechanism.

The experiments were performed on identifiable neurones from the subesophageal ganglia of *Helix aspersa*, usually F1, F2 and E1, E2 according to Kerkut *et al.*⁵. The nerve cell bodies were isolated by dissection and aspirated into a suction pipette for internal perfusion and voltage clamp⁶⁻⁸. A separate intracellular micropipette provided the membrane potential signal for the clamp circuit. K currents were suppressed by tetraethylammonium (TEA) 50 mM extracellularly, 4-aminopyrine (4-AP) 5 mM extracellularly, TEA 10 mM intracellularly, and substitution of K ions by Cs ions intra- and extracellularly. Na currents were suppressed by Tris substitution for Na ions extracellularly and Na ions were absent from the intracellular solution. Intracellular Ca activity was buffered to 10^{-8} M by 0.1-1.0 mM EGTA. In these conditions two currents are present over the potential range of -50 to +50 mV—a linear, non-voltage-dependent, non-time-dependent leakage current and I_{Ca} (ref. 8). Leakage and linear components of capacitative currents were eliminated by adding to a pair of voltage-clamp

*To whom correspondence should be addressed.

Fig. 1 Effects of extracellular ATP (*a*) and AMP-PNP (*b*) on I_{Ca} (current was corrected for linear components of leakage and capacitance by the subtraction method described in text). The compositions of perfusion solutions (in mM) were: internal: Cs-aspartate, 135; TEA-OH, 10; Tris-base, (pH 7.3); external: Tris-Cl, 35; CsCl, 5; TEA-Cl, 50; $CaCl_2$, 10; $MgCl_2$, 15; 4-AP, 5; glucose, 5.5; Tris-base, HEPES, 20 (pH 7.4). All chemicals were added to the control solutions immediately before use. *a*, Upper panel shows the effect of ATP (10^{-6} M) on I_{Ca} elicited by a 70 ms pulse to +20 mV from a holding potential of -50 mV before (○) and 15–20 min after (●) the addition of ATP. Bottom panel shows voltage-current relationships of peak Ca current control (○), 15–20 min after ATP (10^{-6} M) (●) and 60 min after washout (△). *b*, Upper panel shows the effect of AMP-PNP (10^{-8} M) on I_{Ca} with the same pulse protocol as in *a* before (○) and 15 min after (●) the addition of AMP-PNP. Bottom panel shows voltage-current relationships of peak Ca current control (○), 15–20 min after AMP-PNP (10^{-8} M) (●) and 60 min after washout (△).



steps from the holding potential, V_H , of -50 mV the current responses that were equal in amplitude but opposite in sign. At potentials more positive than +50 mV, a time-dependent, nonspecific outward current also appears but these potentials were rarely used in the present experiments and when they were used, nonspecific leakage current was corrected for by subtracting the current that remained after blockage of I_{Ca} by Co substitution for Ca.

The effects of extracellular ATP and AMP-PNP were observed after control currents had been steady for 30–60 min (see Fig. 1). ATP in concentrations of 5×10^{-9} to 10^{-8} M produced an immediate increase in I_{Ca} which continued until a new steady-state level was attained within 5–10 min. Maximum effects were observed at about 10^{-6} M ATP and were similar for Tris, Na and Li salts of ATP, the effects being reversible. AMP-PNP produced the same effect as ATP (Fig. 1*b*). The increases in I_{Ca} amplitude occurred without any apparent changes in the kinetics of activation, and although inactivation occurred at a slightly faster rate for the largest currents, this may be related to Ca-current-dependent inactivation^{8,9}. There were no changes in leakage currents. The increase in Ca currents was not voltage dependent (Fig. 1*a*, lower panel), as the curves of peak current versus voltage before and after ATP were simply scaled versions of each other. This rules out a surface charge effect due to ATP adsorption which would also be unlikely at such low concentrations. As the peak Ca currents are little affected by inactivation, which is at least 10 times slower than activation, and as there is no steady-state inactivation of I_{Ca} at the holding potential of -50 mV used in these experiments, the ATP_o effects must be due mainly to an increase in the channel steady-state activation process. This could result from an increase either in the number of conducting units or in the average unit conductance of channels already present.

Figure 2 shows the dose-response relationship for ATP_o and AMP-PNP_o actions on peak I_{Ca} . The respective average maximal increases were $125 \pm 16\%$ and $130 \pm 16\%$ (mean $\pm$ s.d.) with $n = 6$ for ATP_o and $n = 7$ for AMP-PNP_o. The $K_{0.5}$ is 10^{-6} M for ATP and 3×10^{-9} M for AMP-PNP. At higher concentrations (10^{-3} M) the response falls from its maximum value but we have not further investigated the response at these unphysiological concentrations.

One possible explanation of how ATP_o activates Ca channels is that ATP acts as a substrate for an outward-facing membrane ATPase. However, this would require hydrolysis of the molecule, which is ruled out by the effects of the non-hydrolysable congener, AMP-PNP. Another possibility is an outward-facing

adenylate cyclase, but extracellular cyclic AMP had no effect on I_{Ca} . Alternatively, ATP may permeate these particular neuronal membranes and act as a substrate for adenylate cyclase located in the membrane interior. Investigation of this possibility by perfusing ATP and AMP-PNP intracellularly revealed that ATP produced small increases of I_{Ca} , but required concentrations of $\geq 10^{-5}$ M to achieve the effect, and that AMP-PNP was ineffective even at 10^{-2} M. The effects appeared within 10 min and were steady by 15 min. Hence, the present results are not due to ATP_o acting as a substrate for a membrane system that activates Ca channels. They may be due to a specific reaction between ATP_o and the membrane, as suggested by the purinergic hypothesis, or a mechanism similar to that proposed for the β -adrenergic receptor reaction^{10–12} might exist. In this case the ATP-receptor complex would activate adenylate cyclase to increase cyclic AMP and thus I_{Ca} . Extracellular cyclic AMP had no effect on I_{Ca} but intracellular perfusion at concentrations of 10^{-5} M produced a small increase. This hypothesis would require that the complex amplify adenylate cyclase activity because ATP applied intracellularly in micromolar quantities had much smaller effects than ATP or AMP-PNP applied extracellularly in nanomolar amounts. Extracellular application of ATP might then be enhanced by prior intracellular perfusion with ATP. We tested this proposal using millimolar concentrations of ATP_i and found that further increases of I_{Ca} occurred in response to ATP_o. However, the effects seemed to be simply additive and the results were inconclusive. We therefore tried

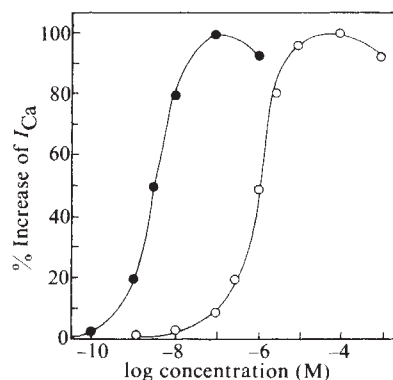


Fig. 2 Dose-response curves for the effects of extracellular ATP (○) and AMP-PNP (●) on I_{Ca} (elicited by a 70 ms long pulse to +20 mV from holding potential of -50 mV). Each point represents the mean of six measurements for ATP and seven for AMP-PNP obtained from different neurones (curves were drawn by eye).

the opposite experiment, to deplete the cell of ATP, and observe the effect of ATP_o. To do this we perfused internally with oligomycin, 10⁻⁶ M (Sigma), or ruthenium red, 5 × 10⁻⁴ M (Sigma). Although the response to extracellular ATP was reduced in these experiments, the Ca currents in the presence of the metabolic inhibitors were also declining, thus preventing any firm conclusion regarding cyclic AMP as a second messenger.

We studied the potency of phosphorylated adenosine and guanosine compounds and found the sequence to be AMP-PNP > ATP > ADP, with AMP and adenosine having no effect. We also found that GTP (Sigma) was about as effective as ADP and that GMP-PNP was as effective as AMP-PNP. Furthermore, the putative purinergic antagonists, apamin (Sigma) 5 × 10⁻¹⁰ M, quinidine (Sigma) 10⁻⁶ M, propranolol (Sigma) 10⁻⁶ and phentolamine (Sigma) 10⁻⁷ M, had no effect on the response to ATP_o. Thus, from the above results the mechanism whereby ATP_o activates Ca channels remains uncertain.

This work was supported in part by DHHS. NIH grants NS-11453 and HL-25145.

Received 28 September 1981; accepted 13 January 1982.

1. Burnstock, G. *Cell Membrane Receptors for Drugs and Hormones: A Multidisciplinary Approach* (eds Smith, R. W. & Bolis, L.) 107–118 (Raven, New York, 1978).
2. Burnstock, G. *J. Physiol., Lond.* **313**, 1–35 (1981).
3. Yatani, A., Tsuda, Y. & Goto, M. *Jap. J. Physiol.* **28**, 47–61 (1978).
4. Flitney, F. W. & Singh, J. *J. Physiol., Lond.* **304**, 21–42 (1980).
5. Kerkut, G. A., Lambert, J. D. C., Guyton, R. J., Loker, J. E. & Walker, R. J. *Comp. Biochem. Physiol.* **A50**, 1–25 (1975).
6. Lee, K. S., Akaike, N. & Brown, A. M. *J. gen. Physiol.* **71**, 489–508 (1978).
7. Lee, K. S., Weeks, T. A., Kao, T. A., Akaike, N. & Brown, A. M. *Nature* **278**, 269–271 (1979).
8. Brown, A. M., Morimoto, K., Tsuda, Y. & Wilson, D. J. *J. Physiol., Lond.* (in the press).
9. Tillotson, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1497–1500 (1979).
10. Tsien, R. W., Giles, W. R. & Greengard, P. *Nature new Biol.* **240**, 181–182 (1972).
11. Reuter, H. *J. Physiol., Lond.* **242**, 429–451 (1974).
12. Rodbell, M. *Nature* **284**, 17–22 (1980).

Human melanoma-associated antigen p97 is structurally and functionally related to transferrin

Joseph P. Brown*, Rodney M. Hewick†, Ingegerd Hellström*, Karl Erik Hellström*, Russell F. Doolittle‡ & William J. Dreyer†

*Program in Tumor Immunology, Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, USA

†Division of Biology 156-29, California Institute of Technology, Pasadena, California 91124, USA

‡Department of Chemistry D-006, University of California, San Diego, La Jolla, California 92093, USA

p97 is a 97,000-molecular weight (MW) cell-surface glycoprotein, which is present in most human melanomas but in only trace amounts in normal tissues^{1–4}. We describe here the purification of p97 by affinity chromatography with monoclonal antibody, followed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and determination of the N-terminal amino acid sequence using a new, highly sensitive protein sequencer⁵. The sequence was found to be homologous to the N-terminal sequences of transferrin and lactotransferrin. This structural homology was confirmed by the observation that antiserum to denatured p97 cross-reacted with denatured transferrin and lactotransferrin. We have also demonstrated that p97 is functionally related to transferrin and lactotransferrin in that it binds iron. This is one of the first reports of the amino acid sequence of a human tumour-associated cell-surface antigen and one of the few cases in which insight has been obtained into its function.

p97 was purified by affinity chromatography. Monoclonal IgG2a antibody specific for p97 was added to a detergent lysate of melanoma cells, and the mixture was passed through a protein A-Sepharose CL-4B column. The antibody and p97 were eluted from the adsorbent at pH 5. The proteins were reduced with 10 mM dithiothreitol (DTT) and alkylated with 50 mM iodoacetamide, and p97 was separated from the IgG heavy and

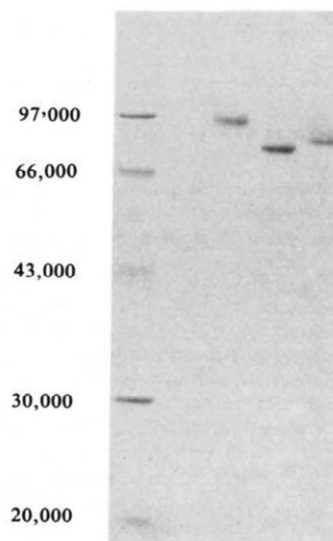


Fig. 1 Purification of p97. SK-MEL 28 melanoma cells (10 g) were suspended in 40 ml 20 mM Tris-HCl buffer pH 8.0, containing 100 mM NaCl, 1 mM EDTA and 0.5% Nonidet-P40 (TNEN buffer) supplemented with 1 mM phenylmethylsulphonyl fluoride (PMSF), and the lysate was centrifuged at 300,000g at 4 °C for 1 h. Antibody 96.5 (5 µg ml⁻¹)⁴ was added, and the lysate was passed through a 1 ml column of protein A-Sepharose CL-4B at 4 °C. The column was washed, then eluted with 2 ml of 100 mM citrate buffer pH 5, containing 0.5% Nonidet-P40. The proteins were precipitated with 60% methanol containing 0.6% acetic acid at -20 °C and electrophoresed on a SDS-6% polyacrylamide gel as described elsewhere¹⁴, except that 5 mM sodium mercaptoacetate was added to the sample buffer. The gel was stained with Coomassie blue, and p97 was eluted by overnight incubation at 20 °C in 0.2 M Tris-acetate buffer pH 8.0, containing 0.1% DTT and 1% SDS, followed by electrophoretic elution in 50 mM Tris-acetate buffer pH 7.8, containing 0.5 mM sodium mercaptoacetate and 0.1% SDS, and then in 50 mM ammonium bicarbonate containing 0.1% SDS. The figure shows a stained SDS-10% polyacrylamide gel of: a, molecular weight markers phosphorylase b (97,000), serum albumin (66,000), ovalbumin (43,000), carbonic anhydrase (30,000) and soybean trypsin inhibitor (20,000); b, purified p97; c, transferrin; d, lactotransferrin.

light chains by preparative SDS-PAGE. Approximately 50 µg of p97 were obtained from 10 g of cells. The purified protein electrophoresed as a single, narrow band on an analytical SDS-polyacrylamide gel (Fig. 1). It appeared to be slightly smaller than phosphorylase b (MW 97,000) and significantly larger than human transferrin and lactotransferrin (MW ~80,000). The immunological specificity of the purification procedure was established using control IgG2a antibodies. As p97 comprised <0.1% of the protein in the initial cell lysate, the two-step procedure resulted in a >1,000-fold purification.

Amino acid sequence analysis⁵ of purified p97 yielded a sequence of 12 residues at the N-terminus (see Fig. 2). A computer search of the *Newat* amino acid sequence library⁶ revealed that this sequence is homologous to the N-terminal sequence of transferrin⁷, seven of the 12 residues being identical (Fig. 2). We consider that this homology is significant for two reasons: first, the two least common amino acid residues, Trp and Cys, are among those conserved, and second, six of the seven conserved residues are also conserved at the N-terminus of lactotransferrin⁸ (Fig. 2). On the basis of the limited amount of sequence data available, p97 appears to be as closely related to transferrin and lactotransferrin (~50% identity) as these proteins are to one another. For comparison, haemoglobin α and β chains have 44% identity⁶. We conclude that p97, transferrin and lactotransferrin have evolved from a common ancestor.

Independent evidence for the structural relationship between p97 and transferrin and lactotransferrin was obtained by

implies a more specialized role for this protein, possibly related to particular aspects of iron metabolism in these tissues.

We thank Cynthia Green and Lynn DeOgny for technical assistance. This work was supported by NIH grants CA 27841, CA 25558, CA 14135, CA 19149 and GM 0695, grant IM 241 from the American Cancer Society and NSF grant PCM 80-05599.

Received 28 October 1981; accepted 5 January 1982.

- Woodbury, R. G., Brown, J. P., Yeh, M.-Y., Hellström, I. & Hellström, K. E. *Proc. natn. Acad. Sci. U.S.A.* **2183**–2187 (1980).
- Brown, J. P., Woodbury, R. G., Hart, C. E., Hellström, I. & Hellström, K. E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 539–543 (1981).
- Woodbury, R. G., Brown, J. P., Loop, S. M., Hellström, K. E. & Hellström, I. *Int. J. Cancer* **27**, 145–149 (1981).
- Brown, J. P., Nishiyama, K., Hellström, I. & Hellström, K. E. *J. Immun.* **127**, 539–546 (1981).
- Hewick, R. M., Hunkapiller, M. W., Hood, L. E. & Dreyer, W. J. *J. Biol. Chem.* **256**, 7990–7997 (1981).
- Doolittle, R. F. *Science* **214**, 149–159 (1981).
- MacGillivray, R. T. A., Mendez, E. & Brew, K. in *Proteins of Iron Metabolism* (eds Brown, E. B. *et al.*) 133–141 (Grune & Stratton, New York, 1977).
- Spik, G. & Mazurier, J. in *Proteins of Iron Metabolism* (eds Brown, E. B. *et al.*) 143–151 (Grune & Stratton, New York, 1977).
- Sutherland, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 4515–4519 (1981).
- Trowbridge, I. S. & Omary, M. B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3039–3043 (1981).
- Enns, C. A., Shindelman, J. E., Tonik, S. E. & Sussman, H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4222–4225 (1981).
- Omary, M. B. & Trowbridge, I. S. *Nature* **286**, 888–891 (1980).
- Trowbridge, I. S. & Lopez, F. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- McConahey, P. J. & Dixon, F. J. *Int. Archs Allergy appl. Immun.* **129**, 185–189 (1966).

Ribosomal RNA transcription *in vitro* is species specific

Ingrid Grummt*, Erika Roth* & Marvin R. Paule†

*Institut für Biochemie der Universität Würzburg, Röntgenring 11, D-8700 Würzburg, FRG

†Department of Biochemistry, Colorado State University, Fort Collins, Colorado 80523, USA

Eukaryotic cells possess three distinct nuclear DNA-dependent RNA polymerases which are responsible for transcription of different sets of genes (for review see refs 1, 2). Recently, cell-free transcription systems have been developed which faithfully initiate transcription of isolated genes by the corresponding RNA polymerase in the presence of crude cellular extracts. These cellular extracts supply additional components required for specific transcription^{3–6}. Successful *in vitro* systems for transcription of RNA polymerase II or III genes were developed using either heterologous or homologous components^{7–11}. In contrast, an analogous cell-free system for the RNA polymerase I transcription unit from mouse has been shown to be active only with homologous extracts from mouse cells⁶. Data presented here show that *in vitro* transcription of ribosomal DNA isolated from mouse, human and a protozoan requires completely homologous components. None of the three active cell-free systems is capable of correct initiation on the nonhomologous templates. Further, supplementation of mouse extracts with purified protozoan RNA polymerase I failed to result in specific transcription of the protozoan rDNA, suggesting that the species specificity of pre-ribosomal RNA synthesis resides, in part, in the transcription factors.

It had been shown previously that specific transcription of rDNA from mouse requires transcription factors from mouse cells; neither human KB cell extracts nor frog oocyte or egg extracts could substitute for extracts from mouse cells⁶. One suggested explanation of this result was that species-specific factors might be involved in the expression of RNA polymerase I genes. We have investigated this apparent species specificity by assaying several different cloned ribosomal RNA genes for their ability to be selectively and correctly transcribed in the mouse extract system. Cloned ribosomal DNAs from *Physarum polycephalum*, *Drosophila melanogaster* and *Xenopus laevis* were used alone or in combination with mouse rDNA as templates in the cell-free transcription system derived from cultured Ehrlich ascites cells. The site of transcription initiation for each of these rDNA genes is known and thus the length of specific *in vitro*

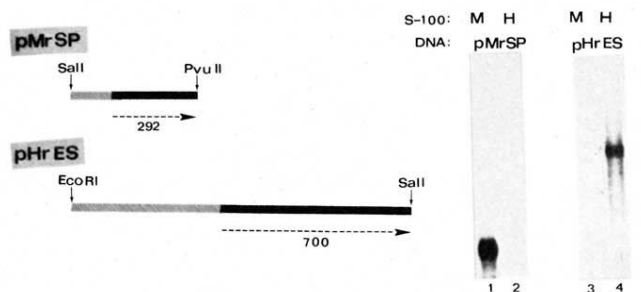


Fig. 1 *In vitro* transcription of rDNAs from mouse and man by homologous and heterologous cell-free transcription systems. One microgramme of cloned mouse rDNA truncated with *PvuII* (pMrSP/*PvuII*, lanes 1, 2) or human rDNA truncated with *SalI* (pHrES/*SalI*, lanes 3, 4) was incubated for 1 h at 30 °C in a 50-μl reaction mixture containing 10 mM HEPES (pH 7.9), 75 mM KCl, 5 mM MgCl₂, 0.05 mM EDTA, 0.5 mM dithiothreitol, 10% glycerol, 10 mM creatine phosphate, 600 μM each of ATP, CTP and UTP, 50 μM GTP, 3 μCi of [α -³²P]GTP (specific activity 10–25 Ci mmol⁻¹) and 25 μl of S-100 extract. S-100 extracts were prepared as described by Weil *et al.*⁴ from cultured Ehrlich ascites tumour cells (EA) or from HeLa cells (HeLa). Synthesis was terminated by addition of 50 μl or 100 mM sodium acetate, 0.5% SDS and 25 μg *Escherichia coli* tRNA. RNAs were extracted with phenol and chloroform, precipitated with ethanol and separated by electrophoresis on a 4% polyacrylamide gel as described by Weil *et al.*⁴. Autoradiograms of the gel are shown. Lane 1: pMrSP/*PvuII* DNA + Ehrlich ascites S-100 extract. Lane 2: pMrSP/*PvuII* DNA + HeLa S-100 extract. Lane 3: pHrES/*SalI* DNA + Ehrlich ascites S-100 extract. Lane 4: pHrES/*SalI* DNA + HeLa S-100 extract.

run-off transcripts can be predicted from restriction analyses and sequence data^{12–15}. None of the heterologous rDNAs yielded specific transcripts (data not shown). The inability of these DNAs to support selective transcription was not due to the presence of inhibitory components in the template preparations because they did not reduce correct transcription of the homologous mouse template in mixing experiments. Thus, although the mouse S-100 extract is capable of directing correct initiation of transcription on mouse ribosomal DNA, it seems incapable of recognizing the promoter for the ribosomal RNA transcription unit from other eukaryotic species.

The above experiments did not rule out the possibility that this apparent species specificity reflects a unique inability for promoter recognition by the mouse transcription system; extracts from other cell types might have a broader capacity to initiate correctly pre-ribosomal RNA chains. To test this, two systems derived from HeLa cells and *Acanthamoeba*, respectively, have been established in addition to the mouse system. In analogy to the mouse system, the human and protozoan S-100 extracts were prepared from logarithmically growing cells and were able specifically to transcribe human rDNA (pHrES) and rDNA from *Acanthamoeba castellanii* (pAr4), respectively (unpublished results). In the experiment shown in Fig. 1, the human and the mouse transcription system was used both with homologous and heterologous rDNA templates and in different combinations of the DNAs and the extracts. The mouse rDNA

Table 1 Specific transcription of rDNA from mouse (pMrSP) and man (pHrES) by homologous and heterologous S-100 extracts

rDNA used as template	Source of extract	Specific transcript*
Mouse	Ehrlich ascites (EA)	Murine
Man	EA	None
Mouse	HeLa	None
Man	HeLa	Human
Mouse	EA + HeLa	Murine
Man	EA + HeLa	Human
Mouse + man	EA	Murine
Mouse + man	HeLa	Human
Mouse + man	EA + HeLa	Murine + human

* The murine transcript is a 292-base long run-off RNA from pMrSP/*PvuII*. The human transcript from pHrES/*SalI* is ~700 bases long.

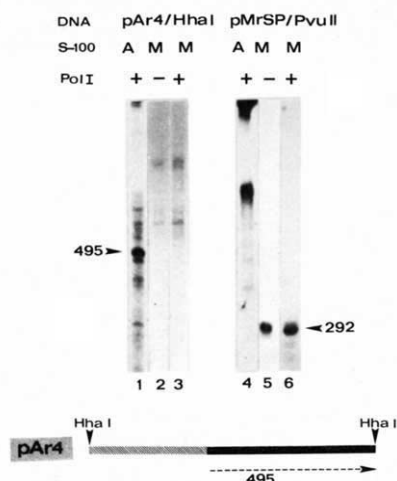


Fig. 2 *In vitro* transcription of rDNAs from *Acanthamoeba* and mouse in the presence of homologous and heterologous S-100 extracts and RNA polymerase I, respectively. A 915-base pair DNA fragment from the *Acanthamoeba* rDNA clone pAr4 (pAr4-HhaI) or a *PvuII* digest of the mouse rDNA clone pMrSP (pMrSP/PvuII) were used as templates and the transcripts analysed as described in Fig. 1 legend. S-100 extracts were obtained either from Ehrlich ascites tumour cells (M) or from log-phase *A. castellanii* cells (A). Reaction mixtures were supplemented (+) or not supplemented (-) with 0.015 U of *Acanthamoeba* RNA polymerase I purified through the heparin-Sepharose step as described elsewhere¹⁷. One unit of RNA polymerase will catalyse the incorporation of 1 nmol of UMP into RNA in 10 min in the conditions described¹⁷. Lane 1: pAr4-HhaI DNA + A S-100 extract + RNA polymerase I. Lane 2: pAr4-HhaI DNA + M S-100 extract - RNA polymerase I. Lane 3: pAr4-HhaI DNA + M S-100 extract + RNA polymerase I. Lane 4: pMrSP/PvuII DNA + A S-100 extract + RNA polymerase I. Lane 5: pMrSP/PvuII DNA + M S-100 extract - RNA polymerase I. Lane 6: pMrSP/PvuII DNA + M S-100 extract + RNA polymerase I.

clone pMrSP, when digested with *PvuII* at a site 292 base pairs downstream from the initiating nucleotide, yields a run-off RNA 292 bases long (ref. 15 and Fig. 1, lane 1). The human clone pHrES yields a transcript ~700 bases long after truncation with *SalI* (E.R. *et al.*, manuscript in preparation, and Fig. 1, lane 4). However, when the mouse DNA (pMrSP/PvuII) is used as template in the human extract system, only low levels of non-specific RNAs are produced (Fig. 1, lane 2). The same holds true if the human template (pHrES/*SalI*) is combined with the mouse S-100 components (Fig. 1, lane 3).

Mixing experiments presented in Table 1 revealed a pronounced template selectivity by the transcription factors for RNA polymerases present in the crude extracts. If both human and mouse rDNA were present in the assay, the source of S-100 determines which template is transcribed. Addition of equal amounts of extracts from human and mouse cells to a mixture of mouse and human templates results in the production of both the 292-base mouse RNA and the 700-base human run-off transcript. These results indicate that the cell-free transcription systems consisting of homologous RNA polymerase I and transcription factor(s) supplied by the S-100 extracts are only able to initiate correctly at homologous promoter sequences.

To verify this conclusion and to discover whether this apparent species specificity resides in the RNA polymerase I or the transcription factors, a third transcription system derived from *A. castellanii* was tested. This system yields a 495-base run-off RNA that coincides with the 5' end of 39S pre-rRNA when a *HhaI*-generated fragment of the *Acanthamoeba* ribosomal RNA repeat unit¹⁶ is transcribed in the presence of S-100 extracts from *Acanthamoeba* (Fig. 2, lane 1). RNA polymerase I purified from this species¹⁷ is incapable of selective transcription of this DNA in the absence of *Acanthamoeba* S-100 extract (Iida and M.R.P., unpublished results). Also, mouse S-100 extract yields no specific transcript with *Acanthamoeba* pAr4-HhaI DNA (Fig. 2, lane 2). When purified *Acanthamoeba* RNA polymerase I is added to mouse S-100 and used to transcribe pAr4-HhaI DNA, an increase in the 915-base 'end-to-end'

transcript but no selective transcription of pre-ribosomal RNA is observed (Fig. 2, lane 3). Supplementation of the mouse S-100 with the protozoan RNA polymerase I also does not stimulate transcription of mouse rDNA (Fig. 2, lane 6); only nonspecific RNAs increase slightly in amount. Thus, supplementation of the nonhomologous transcription factor-template mixture with homologous RNA polymerase I does not result in the synthesis of specific RNA, suggesting that the role of the transcription factor involves recognition of the promoter sequence or the polymerase or of both.

The data presented here strongly suggest that the recognition of rDNA occurs in a species-specific manner. This is in contrast to RNA polymerase II and III transcription systems, which show little or no species specificity. Comparison of nucleotide sequences in front of RNA polymerase II or within RNA polymerase III genes have revealed a remarkable sequence conservation in defined positions necessary for accurate transcription initiation¹⁸⁻²². In contrast, a comparison of the sequence in the region of transcription initiation of different ribosomal genes has revealed no obvious shared consensus sequence among several eukaryotic rDNAs tested^{12-15,23-25}. This lack of a homologous sequence in the rRNA genes is in accord with our finding that species-specific factors are involved in the transcription of rDNA by RNA polymerase I. A change in the single promoter sequence of this transcription unit, if complemented by a similar modification of the polymerase I (or the transcription factor) DNA binding site, could be a reasonably frequent genetic variation. In addition, because the ribosomal RNA transcription unit is repeated many times per cell, genetic experimentation with altered promoter sequences may be a relatively facile (and seldom lethal) event. Apart from the constraints which might arise as a result of the need to keep the repeated genes co-evolving, genetic drift of the rRNA promoter may be comparatively rapid. Conversely, RNA polymerase II deals with thousands of different single copy gene promoters and, because there is probably no direct mechanism to keep them co-evolving, they must remain true to the prototype TATA-box sequence to remain active.

We thank Drs R. Braun, P. Wellauer and R. Reeder for the gifts of rDNAs from *Physarum polycephalum*, *Drosophila melanogaster* and *Xenopus laevis*, respectively, and Dr G. Wilson for the human clone pHrES. This work was supported by grants from the Deutsche Forschungsgemeinschaft (I.G.), the NIH (GM 22580, GM 26059) and by a Senior International Fellowship (F06-TW00524) of the Fogarty International Center (NIH) to M.R.P.

Received 20 November 1981; accepted 6 January 1982.

- Chambon, P. *A. Rev. Biochem.* **44**, 613-638 (1975).
- Paule, M. R. *Trends biochem. Sci.* **6**, 128-131 (1981).
- Wu, G. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2175-2179 (1978).
- Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. *Cell* **18**, 469-484 (1979).
- Manley, J. L., Fire, A., Cano, A., Sharp, P. A. & Gefter, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3855-3859 (1980).
- Grummt, I. *Proc. natn. Acad. Sci. U.S.A.* **78**, 727-731 (1981).
- Waslyk, B., Keding, C., Corden, J., Brison, O. & Chambon, P. *Nature* **285**, 367-373 (1980).
- Proudfoot, N. J., Shander, M. H., Manley, J. L., Gefter, M. L. & Maniatis, T. *Science* **209**, 1329-1336 (1980).
- Sprague, K. U., Larson, D. & Morton, D. *Cell* **22**, 171-178 (1980).
- Dingermann, T. *et al. Nucleic Acids Res.* **9**, 3907-3917 (1981).
- Grummt, I. & Pflugfelder, G. *ICN-UCLA Symp. molec. cell. Biol.* **23**, 303-312 (1981).
- Long, E. O., Rebbert, M. L. & Dawid, J. B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1513-1517 (1981).
- Sollner-Webb, & Reeder, R. H. *Cell* **18**, 485-499 (1979).
- Boseley, P., Moss, T., Mächler, M., Portmann, R. & Birnstiel, M. *Cell* **17**, 19-31 (1979).
- Grummt, I. *Nucleic Acids Res.* **9**, 6093-6102 (1981).
- D'Alessio, J. M., Harris, G. H., Perna, P. J. & Paule, M. R. *Biochemistry* **20**, 3822-3827 (1981).
- Spindler, S. R., Duester, G. L., D'Alessio, J. M. & Paule, M. R. *J. biol. Chem.* **253**, 4669-4675 (1978).
- Goldberg, M. L. thesis, Stanford Univ. (1979).
- Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. *Nucleic Acids Res.* **8**, 127-142 (1980).
- Breathnach, R. & Chambon, P. *A. Rev. Biochem.* **50**, 349-383 (1981).
- Korn, L. J. & Brown, D. D. *Cell* **15**, 1145-1156 (1978).
- Fowles, D. M. & Shenk, T. *Cell* **22**, 405-413 (1980).
- Bayev, A. A. *et al. Nucleic Acids Res.* **8**, 4919-4926 (1980).
- Bach, R., Grummt, I. & Allet, B. *Nucleic Acids Res.* **9**, 1559-1569 (1981).
- Urano, Y., Kominami, R., Mishima, Y. & Muramatsu, M. *Nucleic Acids Res.* **8**, 6043-6058 (1981).

MATTERS ARISING

Lipidic particles and cubic phase lipids

RECENTLY Sen *et al.* have observed¹ regular arrays of inverted micelles in galactolipids and correctly depicted their molecular organization. However, the authors did not differentiate the observed inverted micelles from lipidic particles as seen in phosphatidylethanolamine and cardiolipin²⁻⁵. The inverted micelles observed by Sen *et al.* are, in fact, units of an inverted cubic phase first proposed by Luzzati and Reiss-Husson⁶. The inverted cubic structure has recently been observed in glucolipids by proton NMR⁷ and in monoglycerides by X-ray diffraction^{8,9}. The dimensions reported by Sen *et al.* in galactolipids (8–9 nm) are not too different from those measured by X-ray diffraction in sunflower monoglyceride⁸ (5.8 nm) and in synthetic monoglycerides⁹ (4.5 nm). Unfortunately, Sen *et al.* did not recognize the cubic structure seen in their electron micrographs, and they did not correlate their findings with existing X-ray diffraction and NMR data of similar lipids. Instead, they compared

those spherical inverted micelles with the conical 'lipid particles' seen in phospholipids²⁻⁵.

It is important to note that the 'particles' shown in ref. 1 are of a different nature from the 'lipid particles' discussed previously^{2-5,10-12}. We illustrate this by showing (Fig. 1) a rapidly quenched freeze-fracture (without cryoprotectant) micrograph of mixed egg phosphatidylcholine and dilinoleoylphosphatidylethanolamine (mol ratio 15:85) at 24 °C after being heated to 40 °C. The conical lipidic particles observed in area A are the usual type of bilayer attachment cusps¹² observed in phosphatidylethanolamine-containing lipid mixtures. In area B, the cusps are arranged in a square array, trapping pockets of water between bilayers, thereby forming a complementary array of spheres.

These water-containing spheres, as units of an inverted cubic phase, would be truly inverted micelles were it not for their possible interconnecting water channels^{7,8}. They should not be confused with the lipidic particles, which are the inter-bilayer attachment sites^{5,10-12} surrounding these inverted micelles. The size of the spheres is 17.8 nm in our micrograph, corresponding to our X-ray diffraction data, which show the (200), (220), (222), (422) and (440) diffractions at 8.7, 6.2, 5.1, 3.6 and 3.2 nm of a cubic structure (Fig. 1, inset). Our model for the involvement of lipidic particles in the bilayer-inverted cubic transition will be presented elsewhere.

In conclusion, we have shown that the inverted cubic phase can be observed by electron microscopy, and that this structure occurs not only in glycolipids but also in fully hydrated phospholipids. The units of the inverted cubic phase are inverted micelles, which are distinguishable in shape and size from the so-called lipidic particles or cusps.

S. W. HUI
L. T. BONI

Roswell Park Memorial Institute,
Buffalo, New York 14263, USA



Fig. 1 Freeze-fracture electron micrograph of mixed dilinoleoylphosphatidylethanolamine and egg phosphatidylcholine (mol ratio 85:15) dispersion. The sample was heated to 40 °C and cooled to 24 °C before being freeze-quenched in liquid propane without cryoprotectant, using the quick-freeze sandwich method¹³. Conical pits (lipidic particles) in area A are randomly distributed but become regularly distributed in area B, and spherical units (shadowed opposite to the pits) of a cubic structure are seen between these pits. Small-angle X-ray diffraction was taken from the same dispersion. The five reflections indicated by arrows are identified in the text. Scale bar, 100 nm.

1. Sen, A., Williams, W. P., Brain, A. P. R., Dickens, M. J. & Quinn, P. J. *Nature* **293**, 488–490 (1981).
2. Vail, W. J. & Stollery, J. G. *Biochim. biophys. Acta* **551**, 74–84 (1979).
3. Verkleij, A. J., Mommers, C., Leunissen-Bijvelt, J. & Ververgaert, P. J. J. Th. *Nature* **279**, 162–163 (1979).
4. Hui, S. W., Stewart, T. P., Yeagle, P. L. & Albert, A. D. *Archs Biochem. Biophys.* **207**, 227–240 (1981).
5. Miller, R. G. *Nature* **287**, 166–167 (1980).
6. Luzzati, V. & Reiss-Husson, F. *Nature* **210**, 1351–1352 (1966).
7. Wieslander, A., Rilfors, L., Johansson, B. A. & Linblom, G. *Biochemistry* **20**, 730–735 (1981).
8. Larsson, K., Fontell, K. & Krog, N. *Chem. Phys. Lipids* **27**, 321–328 (1980).
9. Lindblom, G., Larsson, K., Johansson, L., Fontell, K. & Forsen, S. *J. Am. chem. Soc.* **101**, 5465–5470 (1979).
10. Hui, S. W. & Stewart, T. P. *Nature* **290**, 427–428 (1981).
11. Rand, R. P., Reese, T. S. & Miller, R. G. *Nature* **293**, 237–238 (1981).
12. Hui, S. W., Stewart, T. P., Boni, L. T. & Yeagle, P. L. *Science* **212**, 921–923 (1981).
13. Costello, M. J. & Corless, J. M. *J. Microsc.* **112**, 17 (1978).

SEN *ET AL.* REPLY—Hui and Boni are correct in pointing out that the term 'lipidic particle' has been used in the literature to describe a number of different structures. Unlike these authors, however, we believe these structures to be closely related.

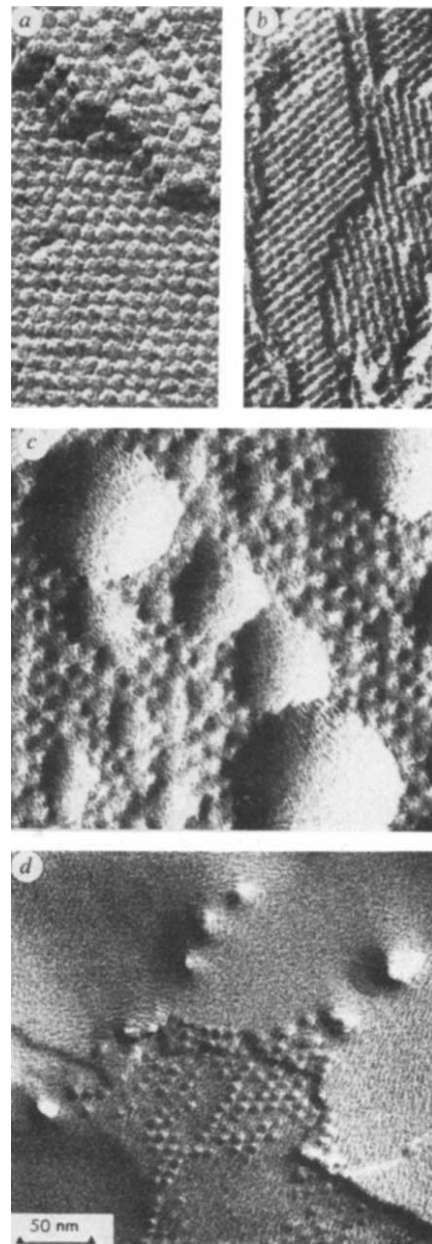


Fig. 2 Electron micrographs of freeze-fracture replicas prepared from sonicated dispersions of mono- and digalactosyldiacylglycerols mixed in a molar ratio of 2:1. *a*, Sample containing 58.4% ethylene glycol; *b*, sample containing no cryoprotectant thermally quenched from 50 °C; *c*, 58.4% ethylene glycol sample; *d*, sample containing no cryoprotectant quenched from ~20 °C. Samples were prepared and replicated as described elsewhere⁵.

We have been able to identify at least four distinct types of lipidic particle from electron micrographs of freeze-fracture replicas of aqueous dispersions of mixed galactolipids. Figure 2 provides a composite view of some of these. The particles shown in Fig. 2a and b are regular quasi-crystalline arrangements induced by the addition of ethylene glycol and by heat, respectively. Figure 2c shows isolated pits formed from particles of the type seen in Fig. 2a. A two-dimensional array of particles located within a single fracture-face is shown in Fig. 2d together with a number of rather larger particles that resemble structures attributed by some workers to inter-bilayer attachment sites¹⁻⁴.

The quasi-crystalline structures, as we pointed out in our original communication⁵, seem to correspond to aggregates of inverted lipid micelles. Hui and Boni have suggested that such structures correspond to isotropic cubic phases of a type previously reported in X-ray diffraction studies of soaps and detergents⁶ and simple monoglycerides^{7,8}. In our experience, a variety of quasi-crystalline structures characterized by different particle sizes and symmetries can be induced to form in mixed galactolipid dispersions. We are, consequently, less willing than Hui and Boni to commit ourselves to any definite symmetry assignments. We do not exclude the possibility that such structures are related to cubic phases similar to those referred to by these authors but, as we have emphasized elsewhere⁹, more evidence based on other techniques is required before any realistic attempt can be made to analyse the detailed structure of the aggregates we have reported.

Hui and Boni further suggest that the occurrence of cusps formed by groups of pits associated with inverted lipid micelles present in these aggregates are directly related to the structures attributed to inter-bilayer attachment sites¹⁻⁴. This resemblance is in our opinion purely fortuitous and merely reflects the fact that such aggregates contain a continuous three-dimensional hydrocarbon matrix. We believe that the various types of lipidic particles reported in the literature can be accounted for in terms of a series of intermediary states made up of inverted micelles at one extreme and bilayer arrangements in which the lipids forming these micelles have been re-incorporated into the bilayer phase at the other.

This reorganization, as detailed elsewhere¹⁰, involves an appreciable expansion of the surface area of bilayer. In the initial stages, this leads to blebs of the type seen in the fracture-face shown in Fig. 2c, and then to a buckling of the fracture-face as shown in Fig. 2d. This process can account both for the formation of tubular-micelle structures and particle-strings of the type shown in Fig. 1 of our original report⁵ and the characteristic arrangements of ridges and particle-strings reported in other studies.

W. P. WILLIAMS*

A. SEN†

A. P. R. BRAIN‡

P. J. QUINN†

Departments of *Biophysics and

†Biochemistry and the

‡Electron Microscopy Unit,

Chelsea College,

University of London,

London SW3 6LX, UK

M. J. DICKENS

MRC Cell Biophysics Unit,

King's College, University of London,

26 Drury Lane,

London WC2R 2LS, UK

1. Miller, R. G. *Nature* **287**, 166-167 (1980).
2. Hui, S. W. & Stewart, T. P. *Nature* **290**, 427-428 (1981).
3. Rand, R. P., Reese, T. S. & Miller, R. G. *Nature* **293**, 237-238 (1981).
4. Hui, S. W., Stewart, T. P., Boni, L. T. & Yeagle, P. L. *Science* **212**, 921-923 (1981).
5. Sen, A., Williams, W. P., Brain, A. P. R., Dickens, M. J. & Quinn, P. J. *Nature* **293**, 488-490 (1981).
6. Luzzati, V. & Reiss-Husson, F. *Nature* **210**, 1351-1352 (1966).
7. Lindblom, G., Larsson, K., Johansson, L., Fontell, K. & Forsen, S. *J. Am. chem. Soc.* **101**, 5465-5470 (1979).
8. Larsson, K., Fontell, K. & Krog, N. *Chem. Phys. Lipids* **27**, 321-328 (1980).
9. Sen, A., Williams, W. P., Brain, A. P. R. & Quinn, P. J. *Biochim. biophys. Acta* (in the press).
10. Sen, A., Williams, W. P., Brain, A. P. R. & Quinn, P. J. *Biochim. biophys. Acta* (in the press).

Gene duplication and the birthday problem

EDLUND and Normark¹ have presented sequence and distribution data of tandem duplications of the *ampC* gene of *Escherichia coli* K-12, which suggest that unequal recombination between homologous sequences of the order of 12 base pairs (bp) in length may be a general mechanism for the generation of tandem duplications. In support of this conclusion they present calculations showing the probability of occurrence of homologies 10-14 bp long in a DNA sequence of given length. For example they state¹ that "on average two homologous sequences of any composition will be found within a random DNA sequence of 4 kilobases (kb)".

It is more realistic to calculate the length of DNA required such that the expected or average number of homologies is one. The number of homologies will be expected to be approximately Poisson distributed and the expected number of homologies, λ , will be given by²

$$\lambda = n \frac{e^{-r/n}}{k!} \left(\frac{r}{n}\right)^k \quad (1)$$

where r is the length of the DNA segment, n is the number of possible different sequences of length m , equal to 4^{12} for a sequence 12 bp long, and k is the number of copies of a homology. Thus, on the average, one pair of homologous sequences 12 bp long will occur in a DNA sequence of length 5.8 kb rather than 4 kb as stated by Edlund and Normark. Allowing for the 1 kb sequence of the *ampC* gene, this value becomes 6.8 kb. The cor-

rected values for Table 1 in ref. 1 may be obtained easily from equation (1).

Table 1 (ref. 1) does not make clear that the probabilities of pairs of homologous sequences are being calculated, rather than the total number of sequences which are homologous in a DNA segment of given length. Although the size of DNA required to obtain an average of four homologous pairs of length 12 bp is 11.6 kb (not allowing for the *ampC* gene), the length of DNA required to obtain an average of one sequence occurring four times is 580 kb! This point is significant as the opportunity for unequal recombination is much greater if multiple copies of a sequence are present than if several pairs of different sequences are present.

A further complication is that the calculation of the probability of occurrence of a sequence homology assumes that the DNA segment of length r contains $r-m+1$ (approximately r) independent sequences of length m . This is clearly incorrect as neighbouring sequences overlap and their probabilities of occurrence will depend on each other. If this dependency is ignored, then the frequency of sequence homologies will be overestimated.

These calculations do not alter Edlund and Normark's conclusion that unequal recombination between random homologies could be a mechanism for generating tandem duplications of the size they observe. This size, 9.8 kb, is still larger than the estimate of the minimum length of DNA required for a homology of length 12 bp to occur an average of once, even if this estimate is biased. However, the importance of such random homologies for generating tandem duplications significantly shorter in length may be questionable.

JULIAN ADAMS

Division of Biological Sciences,
The University of Michigan,
Ann Arbor,
Michigan 48109, USA

1. Edlund, T. & Normark, S. *Nature* **292**, 269-271 (1981).
2. Feller, W. *An Introduction to Probability Theory and its Applications* 3rd edn (Wiley, New York, 1968).

NORMARK REPLIES—Edlund and I recently suggested¹ that unequal recombination between perfect homologies of the order of 10-14 bp may be a general mechanism for generating duplications in the size range of 10 kb. This suggestion was based on DNA sequencing of one duplication in the *ampC* region of the *E. coli* chromosome, on end point determinations of a number of duplications and on probability calculations of occurrence of homologies.

Our calculations that are questioned by Adams must be regarded as approximations. They were made in order to see whether or not our distribution of duplications could be explained by recombination between short homologies of any base composition. Our calculations

and Adams's do not give very different results. The main drawback with both calculations is that we do not know whether or not short homologies are distributed at random in *E. coli* DNA.

What is needed are computer scans for short homologies in continuous stretches of sequenced *E. coli* DNA. At present >5 kb has been sequenced in the *ampC* region of the *E. coli* chromosome (Yaurin and Grunchstrom, personal communication). This region is, however, still too short for such an analysis. Moreover, the recombination site for a number of independent duplications has to be DNA sequenced. Not until then will it be possible to test the calculations made by Adams and our group.

STAFFAN NORMARK

Department of Medical Microbiology,
Stanford University School
of Medicine,
Stanford, California 94305, USA

1. Edlund, T. & Normark, S. *Nature* **292**, 269–271 (1981).

The Red Sea line and Arabian–Nubian magmatism

HARRIS AND GASS¹ suggested that the Red Sea rift developed along the line of a late Proterozoic suture and that compositional differences in the upper mantle and/or lower crust on either side of the suture controlled magmatism within the Arabian–Nubian shields during the Phanerozoic. We now discuss both the existence of the proposed late Proterozoic suture and the nature and distribution of granitic magmatism in the region.

The hypothesis that the present Red Sea rift was controlled by a late Proterozoic suture is unjustified, because it can be demonstrated that there is structural and lithological continuity from the Nubian shield segment across the Red Sea rift zone to the Arabian shield segment.

Figure 1 shows a reconstruction² of the Arabian–Nubian shield before Red Sea rifting. South of latitude 20°N (all localities identified by coordinates are on the Saudi Arabian side of Fig. 1) lineaments are oriented N–S, that is, angled at between 20° and 90° to the trace of the proto Red Sea fracture. There is good lithological correspondence between rocks assigned to the Mormora, Adola, Tambian and Tsaliel stratigraphical units in Ethiopia³ and rocks assigned to the Baish, Bahah, Hali and Jiddah units in Saudi Arabia^{4–6}. Radiometric age data from Saudi Arabia indicate that this distinctive lithological belt was formed between >900 and ~600 Myr ago⁷.

Between latitudes 20° and 24° N, north-east structural trends predominate and the basement fabric is approximately normal to the proto Red Sea fracture. Major faults and distinctive folds can be traced from north-east Ethiopia and north-east Sudan into the west-central Arabian

shield, and lithological correlations can be made between stratigraphical units such as the Nafirdeib (>750 Myr), Asoteriba (~660 Myr), Awat and Homogar of Sudan⁸ and the Samran (>760 Myr) and Fatimah (~680 Myr) of Saudi Arabia⁹.

North of latitude 24°30' N the predominant trend in the basement is north-west and in a few places, for example, between latitudes 26° and 27° N, the basement fabric is consistent with the trace of the proto Red Sea fracture. In this region the Farri (>900 Myr), Al Ays (725–800 Myr) and Hadiyah (~670 Myr) units of Saudi Arabia¹⁰ can be lithologically matched with the Abu Ziran, Dokhan and Hammamat units of south-east Egypt^{11,12}.

The strong north-west trending basement fabric persists north of 27° N within the corridor produced by the Najd strike-slip fault system. This system traverses the proto Red Sea fracture into the Egyptian segment of the Nubian shield. In this region the Zwam Nusfat and Thalbah-Hadiyah-Shammar units of north-west Saudi Arabia¹³ are possibly equivalent to the Abu Ziran, Dokhan-Hammamat units of the central Eastern desert of Egypt.

Whilst basement structures may have locally influenced the development of the Red Sea rift it is highly improbable that the rift was fundamentally controlled by a late Proterozoic suture. Such a cryptic suture would have to traverse at least three major structural provinces in which either north-west, north-east or north-south trending structures are dominant. Furthermore, the remarkable similarity between late Proterozoic stratigraphies of countries bordering the Red Sea is incontrovertible evidence that such a suture did not exist.

Harris and Gass propose that the late Precambrian calc-alkaline granites of the Arabian shield originated in a wet mantle above a subduction zone, whereas, the peralkaline granites resulted from partial melting of dry mantle after subduction had ceased. This hypothesis does not explain why (1) in the area east and south-east of Yanbu Al Bahr, near the central Red Sea coast in the Saudi Arabian shield, peralkaline granitic plutonism occurred some 50–100 Myr before a major episode of calc-alkaline granitic plutonism (D.B.S. and J. Aleinikoff, work in progress), and (2) in general, voluminous calc-alkaline and peralkaline granitic magmas were emplaced in the northern half of the Saudi Arabian shield contemporaneously either in the same pluton^{14,15} or in the same region^{1,10,14,16}.

Finally we discuss whether the absence of undersaturated Phanerozoic magmas in the Arabian peninsula is due to fundamental differences in the chemistry of the upper mantle underlying the Nubian and Arabian shields¹. Whilst there is an apparent paucity of undersaturated Phanerozoic igneous rocks in the Arabian peninsula it is premature to say they are

absent. The undated feldspathoid bearing Sawda syenite pluton in the Midyan area¹⁷ is considered to be a possible correlative of the Egyptian nepheline syenites, and

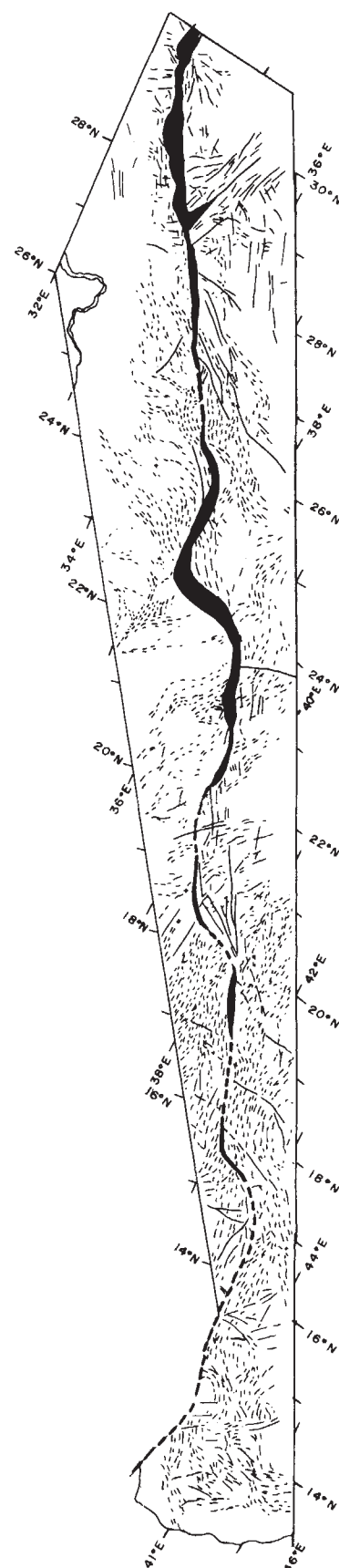


Fig. 1 Arabian–Nubian shield before Red Sea rifting.

peralkaline trachytes of unknown age puncture Palaeozoic rocks in the Tabuk area¹⁴. The chemical similarity of Cenozoic basalts in the Afar, Ethiopia¹⁸, and the Arabian peninsula¹⁹ indicates that the mantle source region was not fundamentally different beneath the Arabian and Nubian shields.

As the Phanerozoic undersaturated alkaline magmatism in the northern Nubian shield has different character from that of the late Proterozoic Pan-African granitic magmatism, we see no reason to assume that this magmatism is an extension of the Precambrian episode. It would be more objective to say that the entire Arabian-Nubian shield experienced late Proterozoic calc-alkaline to peralkaline magmatism, and that a restricted province of predominantly undersaturated to peralkaline magmatism developed in the northern Nubian (and possibly north Arabian) shield region. Phanerozoic alkaline and peralkaline magmatism is found elsewhere in north Africa and that the Phanerozoic Nubian magmatism might be related to events in these regions.

We thank A. R. Drysdall for comments.

N. J. JACKSON

*Saudi Arabian,
Directorate General of Mineral Resources,
Jeddah, Saudi Arabia*

D. B. STOESER

*United States Geological Survey
Saudi Arabian Project, Jeddah,
Saudi Arabia*

- Harris, N. B. W. & Gass, I. G. *Nature* **289**, 395–396 (1981).
- Greenwood, W. R. & Anderson, R. E. *U.S. geol. Surv. Saudi Arabian Proj. Rep. No. 197* (1975).
- Kazmin, V., Shifferaw, A. & Balcha, T. *Geol. Rdsh.* **67**, 531–546 (1978).
- Brown, G. F. & Jackson, R. O. *Bull. Inst. appl. Geol. (Jeddah)*, **1**, 3–9 (1979).
- Jackson, N. J. & Ramsay, C. R. *J. geol. Soc. Lond.* **137**, 617–634 (1980).
- Jackson, N. J. *J. geol. Soc. Lond.* **137**, 629–634 (1980).
- Fleck, R. J., Greenwood, W. R., Hadley, D. G., Anderson, R. F. & Schmidt, D. G. *U.S. geol. Surv. Saudi Arabian Proj. Rep. No. 245* (1979).
- Vail, J. R. *Bull. Overseas Geol. Miner. Resour.* **49**, (1978).
- Gilboy, C. F. & Skiba, W. *Saudi Arabian Dir. Gen. Mineral Resour. Map GM-35* (1979).
- Kemp, J., Pellaton, C. & Calvez, J. Y. *Open-File Rep. OF-01-1*, (Bureau de Recherches Géologie, Miniers, 1981).
- Akaad, M. K. & Noweir, A. M. *Bull. Inst. appl. Geol. (Jeddah)* **4**, 127–135 (1980).
- El-Ramly, M. F. *Ann. Geol. Surv. Egypt* **2**, 1–18 (1972).
- Davies, F. B. *D.G.M.R. Open-file Rep. 732* (1979).
- Stoeser, D. B. & Elliott, J. E. *Bull. Inst. appl. Geol. (Jeddah)* **4**, 1–28 (1980).
- Harris, N. B. W. & Marriner, G. F. *Lithos* **13**, 325–337 (1980).
- Baubron, J. C., Delfour, J. & Vialette, Y. *Open-file Rep. 76 JED 22* (Bureau de Recherches Géologie, Miniers, 1976).
- Ramsay, C. R. *et al. D.G.M.R. Open-file Rep. DGMR-OFR-03* (1981).
- Barberi, F. *et al. J. Geol.* **80**, 720–70 (1972).
- Coleman, R. G., Fleck, R. J., Hedge, C. E. & Ghent, E. D. *Bull. Saudi Arabian Dir. Gen. Miner. Resour.* **22**, D1–D30 (1977).

HARRIS AND GASS REPLY—Jackson and Stoeser have not distinguished between the identification of a Phanerozoic alkaline province in north-east Africa and our speculations on possible causes of this event and its apparent absence in Arabia.

Jackson and Stoeser identify an undated syenite and trachytes from north Arabia to counter our assertion that the Red Sea marks the eastern boundary of the Phanerozoic alkaline province. Future research may discover instances of pre-30 Myr Phanerozoic magmatism in the Arabian shield, but previous work has shown that the plethora of such intrusions in north-east Africa is not present in Arabia. Current geochronological studies on samples from both sides of the Red Sea substantiate the assertion that this igneous province represents the most significant geological contrast between the two sides of the Red Sea.

Following the identification of this igneous province we gave three reasons for its eastern margin being coincident with the Red Sea whilst emphasizing that more data are needed before we can be sure which of these is most relevant. Most of Jackson and Stoeser's comments are directed against our preference for the subduction zone model.

Many have commented on the apparent continuity of structures across the Red Sea, but with coastal plains covered with late Tertiary sediments and rotation of Arabia during the opening of the Red Sea the best that can be postulated is that there seems to be, in many places, structural continuity. The lithostratigraphical correlation between Arabia, Ethiopia, Sudan and Egypt has been shown to be unconvincing even within the southern Arabia shield¹, and correlation of ill-defined and regionally unknown stratigraphical units in the northern Arabian shield is highly questionable. To extend this 'correlation' across the Red Sea with no geochronological or geochemical control is misleading. For example, results from geochronological studies from several of the Egyptian units mentioned by Jackson and Stoeser are totally at variance with their 'lithological matching' of Arabian units. The conclusion that there is precise correlation through the Proterozoic across the Red Sea which precludes any relative movement between north-east Africa and Arabia during late Proterozoic times cannot be substantiated by available data. However, our prime purpose was to identify the Phanerozoic alkaline province in north-east Africa. We then argued that a post-500 Myr distinction in upper mantle geochemistry across the Red Sea line seemed likely. Finally we speculated that an eastward dipping subduction zone with a surface expression along the line of the Red Sea could explain the presence of alkaline magmatism in Africa and its seeming absence in Arabia. If future studies preclude such a suture, this will indicate that the Phanerozoic alkaline province in north-east Africa is the only crustal expression of the geochemical difference in the upper mantle across the Red Sea line until the opening of the Red Sea at 30 Myr.

Jackson and Stoeser imply that the peralkaline magmatism of Arabia is part of the calc-alkaline event, and requires no change in tectonic setting. Whilst the switch from calc-alkaline to peralkaline magmatism was diachronous, the appearance of metaluminous granitoids within peralkaline complexes does not support their view as such granitoids have 'within plate' trace element characteristics (such as high Nb) and are quite distinct from the widespread homogeneous calc-alkaline plutons². Such a marked change in high field-strength element abundance requires a variation in the volatile composition of the upper mantle which could be due to a change in tectonic setting. There are several granitic provinces, such as that of Mesozoic age in Nigeria³, which lie in a within plate tectonic setting and include peralkaline and metaluminous types both with characteristic high field-strength trace-element abundances. It seems realistic therefore to use the terms 'within plate' and 'subduction zone' based on diagnostic trace element abundances, rather than the now outdated calc-alkali-peralkali classification, to identify granitoid provinces.

We believe that the contrast in Phanerozoic magmatism in north-east Africa and Arabia is well-established and the points raised by Jackson and Stoeser do not challenge this. Its relationship with earlier subduction processes suggests that contrasting mantle geochemistry may prove to be the reason and, if so, the Red Sea line was defined in the mantle (and possibly in the crust) as early as 500 Myr.

We thank R. M. Shackleton, A. C. Ries and H. J. Duyverman for comments.

N. B. W. HARRIS

I. G. GASS

*Department of Earth Sciences,
The Open University,
Milton Keynes MK7 6AA, UK*

- Greenwood, W. R., Hadley, D. G., Anderson, R. E., Fleck, R. J. & Schmidt, D. L. *Phil. Trans. R. Soc. A280*, 517–526 (1976).
- Harris, N. B. W. *Tectonophysics* (in the press).
- Bowden, P. & Turner, D. C. in *The Alkaline Rocks* (ed. Sorensen, H.) 330–354 (Wiley, New York, 1974).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

BOOK REVIEWS

Gerontology in adolescence

Alex Comfort

THE most promising omen for experimental gerontology is that it has emerged from the infancy-period of all-explaining general theories into that of specifics. The exercise is to understand and control age changes: these, in a complex organism (human beings would be the obvious beneficiaries — nobody needs long-lived mice), are bound to be multifarious. The first target is to ascertain which time-linked changes are lifespan determining, and the second to establish where, at the supra-genetic level, these are coordinated. Co-ordinated they must be, in view of the high stability of specific lifespans and their divergence in related species.

The first two volumes of this report of a symposium convened by L'Institut de la Vie deal respectively with biology and with medicine and social science. (Volume 3, *Behavioral Sciences and Conclusions*, will be published later this year.) The first, gerontological, volume is an excellent introduction to the diversity of current research topics and the state-of-the-art in each. In contrast to older, one-process views of ageing, the growth stocks in current age studies are immunology, endocrinology, non-actuarial measures of the ageing process — essential, if we hope to look for interventions which modify it in organisms whose lifespans are of the same order as that of the investigator — and, over-arching these, the possibility that the peripheral changes are modulated by a neuroendocrine timekeeper. The odds are that there are some clocks running which are coordinated in this way, and others, possibly including the finitude of cellular clonal division (reviewed in this book by its discoverer Hayflick) which are autonomous.

Two excellent reviews of cerebral ageing, by Scheibel and Scheibel, and by Hachinski and Obrist, open the symposium. Immunology is reviewed by Walford, and hormones by Bierman, Adelman and Everitt — all distinguished investigators. What is lacking, though, is an integrative-speculative model. The discussion of the chances of lifespan modification, opened by the late George Sacher, is largely taken up with the potential of genetic engineering, to the exclusion of the other, and more accessible, programme store, the hypothalamus; and there is little or no consideration of the one authentic instance of age-retardation, that produced by calorie restriction in rodents, where the effect is almost certainly brain-mediated; much of the endocrinological evidence cited by

Ageing: A Challenge to Science and Society. Vol. I *Biology*, edited by D. Danon *et al.*, pp.346, ISBN 0-19-261254-9; Vol. II *Medicine and Social Sciences*, edited by A.J.A. Gilmore, W.M. Beattie *et al.*, pp.403, ISBN 0-19-261255-7. (Oxford University Press: 1981.) Vol. I £25, \$59.50; Vol. II £30, \$59.50.

Adelman and Everitt points in this direction. The hypothalamus would certainly be an easier point of access than the genome, if we are undeterred by Sacher's pessimism over the social effects of success resulting in "demographic catastrophe".

That pessimism, in fact, assorts ill with the urgency of the problems raised by old age in the social and medical areas, the subject of Vol. II. Notably missing from both volumes is a full discussion of senile dementia, which, though a disease, not an "ageing effect", is now a socially ruinous epidemic and a subject of some of the most distinguished current research. There is, however, an excellent anatomical discussion by Terry in the biological section.

Perusal of the volumes leads to several conclusions — that gerontology has advanced enormously, from fringe-discipline to major area of research; that in contrast to its teething years it now has more bricklayers than architects; and that possibly some of the researchers are lured by the prospects of success into long-term, genetic-engineering solutions, when they should be looking for answers nearer home which might have a more immediate clinical application. Clinical applications, not sci-fi projects aimed at radical lifespan change, are the most likely mode of entry for any results of research, and this entry cannot be stopped by social considerations, which may in any case be too alarmist.

Sacher, for example, concludes that we should forgo our arguments about the desirability of life extension until we have done the kind of biologically-based actuarial and socioeconomic modelling that would give us at least a rough idea about the actuarial consequences of . . . life table alteration.

The hazard he envisages here — longer infirm and dependent life — seems biologically unlikely. More serious threats seem to follow from the fact that, while Vol. II of the symposium draws an excellent picture of the medical and social needs of the old under present actuarial conditions, society has so far proved resistant to providing these, and its response to longer health and longer working ability might be

equally negative and belated. A zero-population-growth society calls for long-lived adults: having ceased to be like mice and rabbits, we might gain from becoming elephants. But, as Vol. II well illustrates, applied gerontology will undoubtedly be applied, as soon as it can be, to slowing or postponing disease changes — with the difference from current methods that the approach will probably be fundamental, and generate a whole field of rate-determining medicine. It will, after all, probably be easier to postpone malignancy or immune failure than to reverse it once it has occurred. Examples already exist of cases in mice where diseases can be "postponed out" of the lifespan.

This likelihood makes it extremely important to determine which of the deteriorative processes involved in the loss of homeostasis are self-timed, and which are under more general control and therefore potentially accessible *en bloc* — it is easier to change the policy of a company by tackling the chairman than by nobbling the sales representatives. On these grounds alone one may hope that an overall neurochemical "clock" will be identified.

It is, of course, possible, as Hayflick suggests, that relatively soon we may become able to address a higher level still, and make direct postnatal interventions in the genome. That would, probably, raise some policy questions, but those concerned with lifespan would be dwarfed by other and more general issues.

The most significant contributions to the second volume are Brocklehurst's discussion of training in geriatrics — a particularly urgent problem in the USA — and some short papers on geropsychiatry from the biochemical and genetic points of view. The demographic and social papers, apart from providing some menacing forecasts of what will result from continued neglect of the old, are more general, and a few, in the manner of social gerontology, verge on grannyology. Much of these may be summarized in a quotation from a notable letter to *The Sunday Times*:

I am a senior citizen [and] the following events may reasonably occur. (1) I shall have a gang of young thugs sent to paint my kitchen instead of going to prison (2) I shall have patients from the local mental hospital drafted to dig my garden (3) I may be forced to go to suitable entertainments, drink tea and wear a paper hat . . . We . . . are in a terrifying position: we are recipients. Hands off, please! I am in charge of my life.

That brief testimony should be inscribed

on the brows of the growing phalanx of "social gerontologists".

Although these proceedings are not the last word on age studies (the meeting was held four years ago and much has happened since that time) they make an excellent, authoritative and thoughtful review of the field. The papers by Hachinski and Obrist, by Scheibel and Scheibel, and by Ord, on different aspects of brain ageing, merit reprinting where physicians will see them. □

Alex Comfort is Consultant in Geriatric Psychiatry to the Veterans' Administration, Los Angeles, and an Adjunct Professor at the University of California, Los Angeles.

Search for solutions

Barbara W. Low

Structural Studies on Molecules of Biological Interest: A Volume in Honour of Dorothy Hodgkin. Edited by G. Dodson, J.P. Glusker, and D. Sayre. Pp.610. ISBN 0-19-855362-5. (Oxford University Press: 1981.) £39, \$98.

THE arts and pleasures of praise, in writing, as in reading, may be direct or oblique. This *Festschrift*, written in honour of Dorothy Hodgkin, whose seventieth birthday it celebrates, abundantly manifests these several purposes with grace and rich diversity.

There are contributions from friends, colleagues and former students. Among these latter, Dennis Riley evokes the remarkable limitations of the early laboratory in the University Museum at Oxford (almost a gaslight era), and recalls their essential irrelevance. But he remembers, as we all remember, Dorothy Hodgkin's unique gift for proposing exciting problems without obvious means of solution and the exhilaration and triumph of their solution — learning to swim by being dropped in the deep end.

The section on insulin might be the transactions of a symposium, with all the original sense of convivial atmosphere and stimulating discourse. It provides a brief but fascinating account both of early studies and of those which have followed the determination of the insulin structure in Dorothy Hodgkin's laboratory. There is a continuity of theme in the contributions from different disciplines, for all are built around the concept of structure and function. No where is this more evident than in the account by Zhang You-Shang of the Shanghai Insulin Research Group. Zhang describes the investigation, using semi-synthetic methods, of insulin activity associated with modification of the C-terminal region of the B chain. The activity of D-Ala B23 deshexapeptide insulin is neatly accounted for in terms of the Ramachandran angles of B23 Gly

(normal residue) which lie, in the Ramachandran angle plot, within the permitted region for the D-configuration.

From the contribution of the Peking Insulin Research (Structure) Group, we learn that X-ray crystal structure and solution studies of the B chain C-terminal despentapeptide and desheptapeptide insulins are in progress, with their promise of a closer appreciation of structure/function relationships and the relationship between binding site and active site. Mercola and Wollmer provide an illuminating example of the enhanced understanding of solution phenomena which the high-resolution study of a protein may bring. This section also includes a nice discussion by Sakabe and colleagues of hydrogen bonding in the insulin structure at 1.2 Å resolution. These contributions abound in different ways the fundamental question of intrinsic molecular flexibility or rigidity, particularly at the binding site. I also enjoyed the discussion of Falkmer and Emdem on insulin evolution and Steiner's contribution on insulin precursors.

I have chosen to write of this section in the greatest detail because it represents not only the continuity of Dorothy Hodgkin's interests, but also a continuity of theme and focus. Crystallographers will especially enjoy the methods section, which includes a discussion by Bentley and Mason of the neutron diffraction study of monoclinic lysozyme and a contribution by Isaacs on fast-Fourier least-squares refinement. The more general section on protein structure includes an account by Harrison — "Approaches to the Structure and Function of Ferritin" — a history of crystal structure studies of the virus-like iron storage protein apoferritin. She concludes this provocative contribution with a reference to the isolation of a ferritin-like protein from *Azotobacter vinelandii* and from *E. coli*. Vainshtein also spans the range of common function by describing the structure of leghaemoglobin (from *Lupinus* nodule extracts) and comparing it with the functionally-related myoglobin.

There are also rewards elsewhere in the book. The contributions by Glusker on the vitamin B₁₂ work and by Harding on the study of gramicidin S structure both provide a lively sense of the experimentalist tackling difficult problems by persistent frontal attack and by flanking approaches, reconsidering constantly as all evidence is re-evaluated. Thus difficulties may be approached almost by stealth — the coaxing of a problem to its solution.

There are many other stimulating and pleasing contributions. But to choose was here a regrettable necessity. The editors deserve our warmest congratulations and gratitude. □

Barbara W. Low is Professor of Biochemistry at Columbia University. She worked under Dorothy Hodgkin's supervision and her doctoral thesis described the X-ray crystal determination in three dimensions of potassium and rubidium penicillin.

Medicines worldwide

John Butterfield

Pharmaceuticals and Health Policy: International Perspectives on Provisions and Control of Medicines. Edited by R. Blum et al. Pp.272. ISBN 0-7099-0608-0. (Croom Helm/Holmes & Meier: 1981.) £17.95, \$37.50.

THE provision of medicines in the modern world is a complex process. It begins with the research stage; proceeds through toxicity testing and government-regulated clinical trials, drug licensing and registration; then passes into the distribution phase, which includes promotion by the manufacturers, dispensing by pharmacists and, in the case of certain medicines, prescription by the medical profession; finally it ends with consumption by patients. The beliefs and knowledge of the latter may critically affect patterns of actual drug usage in the home, a fact too often ignored by the professionals who direct the previous stages.

The authors of *Pharmaceuticals and Health Policy* collectively espouse the view that improved health care depends, in part, on the improved supply of medicines. To this end they present information about each of the key stages of the provision process and outline the various, sometimes conflicting, interests of the actors in the drama. Wisely, then, the editors have combined the skills of individuals from many disciplines — including medicine, psychology, administration, pharmacology, epidemiology, economics and sociology — and the result is an impressive amalgam of different viewpoints.

In tackling such a broad subject, the volume inevitably has its weaknesses. For example, despite the fact that readers are urged to identify themselves as "friends of research and discovery", little attempt is made to discuss the alternative forms of research management and funding which a community might adopt. Under what conditions is marked-funded, industrial research superior to state-funded, academic research? What acceptable or unacceptable "disbenefits" may be associated with each of these different modes of research support? Why have the international pharmaceutical houses been the source of so much of the pharmacological innovation of recent decades?

Similarly, despite a long essay on the individual use of medicines, no clear image emerges as to how the authors envisage consumers might be influenced to create more appropriate and productive behaviour in the field of drug usage. Personally, I would have preferred to see a more specific consideration of the needs and problems of that section of the population that conventional health education cannot reach, that is the poorer people — the people for whom education may be an anathema, who may be mentally less able

or emotionally less mature, but people who are the consumers likely to be at the greatest risk both of suffering drug-related mishaps and of not receiving suitable medication for treatable conditions. This is where health promotion must now be directed, and vigorously.

In addition, the volume does not adequately examine the consequences that national government policies have on the processes of medicine provision throughout the world. It is increasingly important that political interests are not exercised in areas such as drug registration, for example, where they could pass without the loud critical comment so often directed at academic, professional and commercial bodies.

But these cautions apart, the volume remains a worthwhile achievement. Lumbroso's essay on the introduction of new drugs and the problems faced by clinical trial organizers is of particular value. So too is Catherine Stenzl's concise chapter on the role of international organizations in medicine policy. Herxheimer and Lionel provide a sensible closing contribution on coherent policy formation, outlining the difficulties present and options open both nationally and internationally. It is with respect to the issues surrounding the latter area that the main messages of the book are to be drawn.

In the countries of the developed world, the home ground of pharmacological innovation, we have moved in the past few decades to a delicate balance between rival public, commercial, professional and governmental interests. "Structural maturity" may still not be fully established but the problems that are left are minor compared with those of 30, or even 20, years ago. In the rich world, at least, the contention that we are "reasonably well served by our drug providers" is now an entirely legitimate one.

Yet in the industrially less developed nations the situation is not so satisfactory. Not only are half or two-thirds of the people deprived of modern medication in more or less any form; but those who can obtain drugs do so in a local environment in which they may be open to many forms of harmful practice. These may be a result not only of ignorance, or lack of regulation, but also corruption — commercial, political and professional. In my view this is the central problem which drug providers throughout the world must face up to in the 1980s.

In his preface Dr T.A. Lambo, Deputy Director, WHO, says,

By my lights the reader [of this book] will be every government authority anywhere concerned with drug policy, anyone teaching clinical pharmacology, all concerned with cultural or social medicine, or medical economics, and most certainly my own

colleagues at every level who attend to international issues and collaboration in health affairs.

It certainly covers a very important subject — health policy. Those of us involved like to believe the quality of our work, of the decisions we make, the regulations we frame, the ethical motivation we achieve in the whole pharmaceutical process, are good indices of the quality of the society we seek to serve. The great trick from now on will be not simply to take care of the geese that lay the golden eggs — whether you regard those geese as the minds in the laboratory or the firms who develop and the clinicians who try the new medicines. It must also include ensuring that the eggs are safe for the children playing in the farmyard and — dare one use an old word? — that the charitable impulse of the medical and allied professions will prevail when the eggs get to the market places. All this is a complex network but there is no textbook for those involved. This slim volume will give newcomers valuable check lists based on the surveillance that the distinguished contributors have been maintaining. □

Sir John Butterfield is Regius Professor of Physic at the University of Cambridge.

Great matters

Charles F. Kennel

Cosmic Plasma. By Hannes Alfvén. Pp.164. ISBN 90-277-1151-8. (Reidel:1981.) Dfl.75, \$39.50.

HANNES Alfvén's *Cosmic Plasma* is a highly individualistic essay about many topics central to contemporary space physics, astrophysics and cosmology. It deals with laboratory studies of cosmic plasma processes; the Earth's magnetosphere, ionosphere and aurora; Venus's magnetosphere; the solar wind and heliosphere; comets; the origin of the Solar System; interstellar clouds; cosmic rays; the question of antimatter in the Universe (to paraphrase Fermi: where is it?); quasi-stellar objects; and the expanding Universe. Alfvén's often rebelliously unconventional treatment of these topics leads him to conclude that plasma physics rules the behaviour of the Universe much more than is generally appreciated.

A basic premise of *Cosmic Plasma* is that our knowledge of laboratory and Solar System plasma physics must inform our understanding of more distant plasmas that we can neither manipulate nor probe *in situ*. To Alfvén, it is striking that spacecraft observe very different states of plasma separated by thin, stable current layers. We have learned, he says, that Solar System plasmas, and by implication all

NEW

GENETICS AND PROBABILITY IN ANIMAL BREEDING EXPERIMENTS

EARL L. GREEN

Director Emeritus of The Jackson Laboratory, Bar Harbor, Maine.
Editor, *Biology of the Laboratory Mouse*, Second Edition.

Contents

Probability and Statistics; Segregation of Alleles; Assortment of Non-Alleles; Linkage, Recombination and Mapping; Mating Systems; Appendix 1 — Mean and Variance of a Binomial Distribution; Appendix 2 — Estimation of a Parameter by the Method of Maximum Likelihood; Appendix 3 — Extensions of the Method of Maximum Likelihood; Appendix 4 — Comparative Efficiency of Matings for Detecting and Measuring Linkage; Appendix 5 — the Fibonacci Sequence; Appendix 6 — Systems of Mating; Appendix 7 — Numbers of Mating and Numbers of Mice per Mating; Appendix 8 — Nomenclature; Appendix 9 — Record Keeping; Appendix 10 — Mouseroom Layout and Procedures; Glossary of Signs; Literature Cited; Index.

This book is a primer and reference on probability, segregation, assortment, linkage and mating systems for biomedical scientists who breed and use genetically defined laboratory animals for research.

It is organised so as to be useful to 'novices' and 'experts' alike. The former will find an orderly development of the probability aspects of transmission tests, segregation analysis, allelism tests, independence tests, linkage analysis and mapping. They will see how to design genetic experiments to yield data which can be analysed and they will also find a clear exposition of the theoretical consequences of various breeding systems which will enable them to make rational choices as to the kinds of animals to use in their own research.

Experts will find that the most frequently used formulae needed for estimating various genetic parameters and their random sampling variances are assembled in tables or are clearly set forth in the text and that the extensive table of contents and detailed index will enable them to find the necessary formulae quickly.

LEVEL: Professional, graduate, undergraduate reference.

1981: £20.00; 256pp; ISBN 0 333 27243 9

RELATED TITLES

FESTING: *Inbred Strains in Biomedical Research*. £25.00 1979

FESTING: *Animal Models of Obesity*. £17.50 1979

SPARROW: *Immunodeficient Animals for Cancer Research*. £20.00 1980

Send your orders, requests for further information etc. to: Frances Roach, Scientific and Medical Division, Macmillan Press, Houndmills, Basingstoke, Hampshire, RG21 2XS, U.K.

SCIENTIFIC
& MEDICAL

M

MACMILLAN
PRESS

Erratum: the price of the paperback edition of *Principles of Neural Science* (reviewed in *Nature* 295, 474; 1982) is £19.95, not £25.50.

cosmic plasmas, are cellular or filamentary in structure. From this stems his astonishing conjecture about antimatter. The symmetry between matter and antimatter in the laws of elementary particle physics suggests that the Universe ought to have been created with roughly equal amounts of each. Our Solar System is made of matter, and the lack of evidence for antimatter elsewhere may mean that any primordial antimatter was annihilated when it came into contact with matter, leaving us with a matter Universe. However, Alfvén proposes that cells of matter and antimatter co-exist throughout the Universe, that the cells are kept apart by plasma forces and that annihilation is drastically slowed because it takes place only in the thin layers between cells. Antimatter is difficult to detect except by the radiation coming from its annihilation, which is much less intense than it would be if matter and antimatter were evenly distributed. We do not even know for sure, claims Alfvén, whether our nearest neighbour, α Centauri, is a star or an anti-star. Annihilation, when it does occur, could provide the energy for the most violent cosmic phenomena we observe. For example, the mysterious quasi-stellars might be the result of an occasional star-anti-star collision in the nucleus of an otherwise normal galaxy — if indeed a galaxy made half of antimatter is normal. And annihilation could have provided the energy for the expansion of the Universe. By weaving such (controversial) arguments, Alfvén elevates plasma physics to a position which is central to all of astrophysics.

Alfvén's track record compels us to take *Cosmic Plasma* seriously. He did not win his Nobel Prize for thinking the way others do. Forty years ago, he helped to create the foundations of magnetohydrodynamics. Twenty years ago, when this work had become central to space physics, he was already insisting that the most important effects occurred in thin current-carrying filaments, or sheets that separated uninteresting regions of "passive" plasma. There one would find small regions of strong electric fields parallel to the magnetic field existing in complete violation of magnetohydrodynamic theory. These electric fields were what accelerated particles to high energies in the aurora. Today we know that auroral electrons are accelerated about 5,000 km above the poles of the Earth by electrostatic structures that bear Alfvén's name for them — double layers. He now asserts that the cosmic rays we observe are generated by undetectable double layers far above the poles of the Sun. The absence of antiparticles in the cosmic rays therefore does not argue against antimatter elsewhere in the Universe.

Is all of the physics presented in *Cosmic Plasma* correct? Many readers will find something to disagree with, just as I believe that altogether too much can be

understood assuming that cosmic rays above 1 MeV energy come largely from the Galaxy to accept Alfvén's Solar System origin for them uncritically. Some readers may wish to turn to Eugene Parker's excellent monograph, *Cosmical Magnetic Fields* (Oxford University Press, 1979), for a systematic development of the magnetohydrodynamic point of view Alfvén so thoroughly rejects. Is there another great idea or two lurking in *Cosmic Plasma*? It is too soon to tell. In any case, might the theoretician's cardinal sin be, not to be wrong, but to be uninteresting?

For whom is *Cosmic Plasma* intended? Certainly not the beginning student. It has been said that cosmic plasma theory proceeds from the cartoon approximation to the back-of-envelope approximation, and thence to detailed calculation and synthetic understanding. Alfvén believes that even our cartoons are all wrong, and so his discussion is rich in fresh cartoons, envelope calculations — and his own synthetic understanding — but is seriously lacking in the systematic development students need. One must know the literature thoroughly to follow Alfvén. Although he speaks directly to today's professionals, who can read *Cosmic*

Plasma without a glossary, he fears they won't listen. Maybe he has tomorrow's young researchers in mind.

Alfvén's future biographers will find expressed in *Cosmic Plasma* his intellectual and psychological position *vis à vis* the conduct of science. Historians of science may perceive an interesting parallel. Just as it was given to Newton's generation to comprehend how gravity rules the Solar System, so has it been given to ours to discover and understand Solar System plasmas. Like Newton, Alfvén could not resist the temptation to contemplate the cosmic significance of what he had learned. Perhaps Alfvén's *Cosmic Plasma* is similar in intent to Newton's Letters to Richard Bentley. And the sentiments Newton expressed in the famous *scholium* to the *Principia* are echoed in the final lines of *Cosmic Plasma*:

It seems that so far no phenomena have been discovered which necessarily call for any new laws. The basic properties of a plasma seem to be the same, from the laboratory to the Hubble distance. □

Charles F. Kennel is a Professor of Physics at the University of California, Los Angeles, and a member of its Institute of Geophysics and Planetary Physics.

Ten years after Moon-walking

David W. Hughes

The Moon — Our Sister Planet. By Peter H. Cadogan. Pp.391. Hbk ISBN 0-521-23684-3; pbk ISBN 0-521-28152-0. (Cambridge University Press: 1981.) Hbk £27.50, \$59.95; pbk £12.50, \$24.95.

ON DECEMBER 19, 1972, Eugene A. Cernan and Harrison H. Schmitt left the valley of Taurus-Littrow near the coast of the great frozen basaltic "sea" of Serenitatis. No one has been back to the Moon since. Twelve men wandered 90 km over its surface, spent 80 hours outside their spacecraft and brought back 381 kg of rock and soil.

Ten years on we can sit back and read about what has been learnt. Cadogan wrote his doctoral thesis on the carbon chemistry of lunar samples and worked on lunar dating after that. So we have an author who was at the sharp end of lunar research, taking time to write a book which traces our progress in understanding Earth's sister planet. Cadogan has provided us with a good, general review of the subject; a book which delves into lunar science to some depth and yet assumes no previous knowledge of mathematics or geology. He is to be congratulated on a job well done.

The book is split into six main sections. The first discusses the Moon as seen from a distance, and traces the advances in lunar cartography and geology and also the famous impact versus volcanism debate over crater formation. The second

condenses the enormous amount of information that accrued from the spate of American and Russian landers. Moon rocks and minerals are dealt with in the third section, and the next discusses the lunar soil and the way in which this is affected by the bombardment of micrometeorites, erosion by the solar wind and radiation damage inflicted by cosmic rays. Section 5 looks at the Moon as a whole and studies its shape, gravity, atmosphere, magnetism and heat sources. The final section turns to the problem of lunar origin, age and orbital evolution.

Cadogan's enthusiasm for his subject bubbles from every page, but at the same time he has taken great care to explain difficult concepts. The standard of the figures is excellent and the choice of illustrations and their positioning in the text leave nothing to be desired. The book is worth buying for the figures alone. I have, however, a minor point of criticism. Cadogan's style of English is idiosyncratic to say the least. Why should I be so generous? The style is awful. Even after 200 pages it still shocks one's sensitivity. This is not to say that the book is unreadable — in fact, I enjoyed reading it very much — but still the ghost of the illiterate scientist walks through every paragraph. □

David W. Hughes is a Lecturer in Astronomy and Physics at the University of Sheffield.

BOOKS RECEIVED

Earth Sciences

- BEN-MENAHEN, A. and SINGH, S.J. *Seismic Waves and Sources*. Pp.1108. ISBN 3-540-90506-5. (Springer-Verlag: 1981.) DM182, \$107.40.
- BERGER, A. (ed.). *Climatic Variations and Variability: Facts and Theories*. NATO Advanced Study Institute. 1st Course of the International School of Climatology, Ettore Majorano Centre for Scientific Culture, Erice, Italy, March 1980. Pp.771. ISBN 90-227-1300-6. (Reidel: 1981.) Dfl.175, \$87.50.
- FLETCHER, W.K. *Handbook of Exploration Geochemistry Vol.1, Analytical Methods in Geochemical Prospecting*. Pp.255. ISBN 0-444-41930-6. (Elsevier Scientific: 1981.) \$59.50, Dfl.140.
- HECHT, M.E. and REEVES, R.W. *The Arizona Atlas*. Pp.164. (no ISBN) (University of Arizona, Tucson: 1981.) Pbk np.
- HUSEBYE, E.S. and MYKKELTVEIT, S. (eds). *Identification of Seismic Sources—Earthquake or Underground Explosion*. Proceedings of the NATO Advanced Study Institute held at Voksenåsen, Oslo, Norway, September 1980. NATO Advanced Study Institute Series: Series C—Mathematical and Physical Sciences, Vol.74. Pp.876. ISBN 90-277-1320-0. (Reidel: 1981.) Dfl.195, \$98.
- JENKINS, D.G. and MURRAY, J.W. (eds). *Stratigraphical Atlas of Fossil Foraminifera*. Pp.310. ISBN 0-85312-210-5. (Ellis Horwood: 1981.) £25.
- LEE, W.H.K. and STEWART, S.W. *Principles and Applications of Micro-earthquake Networks*. Advances in Geophysics, Supplement 2. Pp.293. ISBN 0-12-018862-7. (Academic: 1981.) \$32.
- LIPPSON, A.J. *et al.* *Environmental Atlas of the Potomac Estuary*. Pp.279 + 9 maps. (no ISBN) (The John Hopkins University Press: 1981.) £20, \$35.
- MANDEL, S. and SHIFTAN, Z.L. *Groundwater Resources*. Investigation and Development. Pp.269. ISBN 0-12-468040-2. (Academic: 1981.) \$32.
- MOSELEY, F. *Methods in Field Geology*. Pp.211. Hbk ISBN 0-7167-1293-8; pbk ISBN 0-7167-1294-6. (W.H. Freeman: 1981.) Hbk £12; pbk £6.50.
- ROBERTS, A. *Applied Geotechnology. A Text for Students and Engineers on Rock Excavation and Related Topics*. Pp.344. Hbk ISBN 0-08-024015-1; flexi ISBN 0-08-024014. (Pergamon: 1981.) Hbk np; flexi £10.50, \$25.
- SMART, P. and TOVEY, N.K. *Electron Microscopy of Soils and Sediments: Examples*. Pp.177. ISBN 0-19-854515-0. (Clarendon Press/Oxford University Press: 1981.) £30.
- SOUTHWOOD, D.J. (ed.) *ULF Pulsations in the Magnetosphere*. Advances in Earth and Planetary Sciences, 11. Reviews from the Special Sessions on Geomagnetic Pulsations at XVII General Assembly of the International Union for Geodesy and Geophysics, Canberra, December 1979. Supplement Issue to Journal of Geomagnetism and Geoelectricity. Pp.200. ISBN 90-277-1232-8. (Reidel: 1981.) Dfl.60, \$29.50.

Biological Sciences

- ABRAHAMSON, E.W. and OSTROY, S.E. (eds). *Molecular Processes in Vision*. Benchmark Papers in Biochemistry, Vol.3. Pp.431. ISBN 0-87933-372-3. (Academic: 1981.) \$55.
- ADER, R. (ed.). *Psychoneuroimmunology*. Pp.661. ISBN 0-12-043780-5. (Academic: 1981.) \$59.
- AUTRUM, H. *et al.* (eds). *Progress in Sensory Physiology*, Vol. 1. Pp.179. ISBN 3-540-08413-4. (Springer-Verlag: 1981.) DM 69, \$31.40.
- BARBER, H.G. and HAWORTH, E.Y. *A Guide to the Morphology of The Diatom Frustule with a Key to The British Freshwater Genera*. Pp.112. Pbk ISBN 0-900386-24-8. (Freshwater Biological Association, Ambleside, Cumbria: 1981.) Pbk £3.50.
- BATSCHLET, E. *Circular Statistics in Biology*. Pp.372. ISBN 0-12-081050-6. (Academic: 1981.) £28.80, \$69.50.
- BILLINGS, E. and WESTMORE, A. *The Billings Method. Controlling Fertility without Drugs or Devices*. Pp.254. ISBN 0-7139-1454-8. (Allen Lane, London: 1981.) £5.95.
- BISHOP, J.A. and COOK, L.M. (eds). *Genetic Consequences of Man Made Change*. Pp.410. ISBN 0-12-101620-X. (Academic: 1981.) £23.60, \$57.
- CAPLAN, A.L., ENGELHARDT, H.T. Jr. and MCCARTNEY, J.J. (eds). *Concepts of Health and Disease Interdisciplinary Perspectives*. Pp.756. Pbk ISBN 0-201-00973-0. (Addison-Wesley: 1981.) Pbk \$29.50.
- DAVIS, D.D. *The Unique Animal: The Origin, Nature and Consequences of Human Intelligence*, 1. Pp.336. ISBN 0-907152-02-3. (Prytaneum Press, London: 1981.) Hbk £12.95; pbk £6.95.
- DAVISON, A.N. and THOMPSON, R.H.S. (eds). *The Molecular Basis of Neuropathology*. Pp.693. ISBN 0-7131-4374-6. (Edward Arnold, London: 1981.) £65.
- DAY, M.H. (ed.). *Vertebrate Locomotion*. The Proceedings of a Symposium held at The Zoological Society of London, March 1980. Pp.472. ISBN 0-12-613348-4. (Academic: 1981.) £37.20, \$89.50.
- DOWDESWELL, W.H. *The Life of the Meadow Brown*. Pp.165. Flexi ISBN 0-435-60224-1. (Heinemann Educational: 1981.) £5.95.
- DUNCAN, W. (ed.). *Prostate Cancer. Recent Results in Cancer Research*, Vol. 78. Pp.190. ISBN 3-540-10676-6. (Springer-Verlag: 1981.) DM 88, \$40.
- EFFROS, R.M., SCHMID-SCHÖNBEIN, H. and DITZEL, J. (eds). *Microcirculation: Current Physiologic, Medical, and Surgical Concepts*. Pp.315. ISBN 0-12-232560-5. (Academic: 1981.) \$45.
- FINN, C.A. (ed.). *Oxford Reviews of Reproductive Biology*, Vol. 3. 1981. Pp.326. ISBN 0-19-857536-X. (Oxford University Press: 1981.) £30.
- FOSTER, H.L., SMALL, J.D. and FOX, J.G. (eds). *The Mouse in Biomedical Research*. Vol.1, History, Genetics, and Wild Mice. Pp.306. ISBN 0-12-262501-3. (Academic: 1981.) \$54.
- FRIEDMANN, H.C. *Enzymes. Benchmark Papers on Biochemistry*, Vol. 1. Pp.716. ISBN 0-87933-367-7. (Academic: 1981.) \$68.
- GARRATTINI, S. and SAMANIN, R. (eds). *Anorectic Agents: Mechanisms of Action and Tolerance*. Pp.256. ISBN 0-89004-640-9. (Faber & Faber: 1981.) \$28.
- GEORGE, A.S. (Executive ed.) *Flora of Australia*, Vol. 1: Introduction. Pp.200. Hbk ISBN 0-642-06652-3; flexi ISBN 0-642-06653-1. (Australian Government Publishing Service, Canberra: 1981.) Hbk \$12.50; flexi \$9.50.
- HALLER, O. (ed.). *Natural Resistance to Tumors and Viruses. Current Topics in Microbiology and Immunology*, Vol. 92. Pp.128. ISBN 3-540-10732. (Springer-Verlag: 1981.) DM 68, \$31.
- HARRIS, C.L. *Evolution: Genesis and Revelations with Readings from Empedocles to Wilson*. Pp.339. Pbk ISBN 0-87395-487-4. (State University of New York Press: 1981.) Pbk \$9.95; hbk \$29.50.
- HEISER, C.B. Jr. *The Sunflower*. Pp.198. Pbk ISBN 0-8061-17435. (University of Oklahoma Press: 1981.) Pbk \$5.95. (available in hbk).
- HENLE, W. *et al.* (eds). *Current Topics in Microbiology*, Vol. 91. Pp.284. ISBN 3-540-10722-3. (Springer-Verlag: 1981.) DM 120, \$54.60.
- HERSEN, M. and BELLACK, A.S. (eds). *Behavioral Assessment. A Practical Handbook*. 2nd Edn. Pp.603. Hbk ISBN 0-08-025956-1; pbk ISBN 0-08-025955-3. (Pergamon: 1981.) Hbk np; pbk £8.60.
- HESS, H.-J. (ed.). *Annual Reports in Medicinal Chemistry*, Vol.16. Pp.364. Pbk ISBN 0-12-040516-4. (Academic: 1981.) Pbk \$27.50.
- HORNBEIN, T.F. (ed.). *Regulation of Breathing (in two parts)*. Part II: Lung Biology in Health and Disease, Vol. 17. Pp.808. ISBN 0-8247-1013-4. (Marcel Dekker: 1981.) SwFr. 275.
- HRDINA, P.D. and SINGHAL, R.L. (eds). *Neuroendocrine Regulation and Altered Behaviour*. Pp.404. ISBN 0-306-40840-6. (Plenum: 1981.) \$49.50.
- JAMESON, E.W. Jr. *Patterns of Vertebrate Biology*. Pp.177. ISBN 3-540-90520-0. (Springer-Verlag: 1981.) DM 63, \$33.10.
- JANICK, J. *et al.* *Plant Science. An Introduction to World Crops*. 3rd Edn. Pp.868. ISBN 0-7167-1261-X. (W.H. Freeman: 1981.) £15.60.
- JING, Z. and YANGWEN, L. (eds). *The Giant Panda*. Pp.171. ISBN 0-442-20064-1. (Van Nostrand Reinhold: 1981.) £25.45.
- JOHNSON, C.B. *Physiological Processes Limiting Plant Productivity*. Pp.395. ISBN 0-408-10649-2. (Butterworths: 1981.) £25.
- KEYNES, R.D. and AIDLEY, D.J. *Nerve and Muscle*. Pp.163. Hbk 0-521-23945-1; pbk ISBN 0-521-28362-9. (Cambridge University Press: 1981.) Hbk £15; pbk £4.95.
- SCHULTZ, S.G. (ed.). *Ion Transport by Epithelia*. Society of General Physiologists Series, Vol.36. Pp.287. ISBN 0-89004-610-7. (Raven: 1981.) \$32.
- SHAVELSON, R.J. *Statistical Reasoning for the Behavioral Sciences*. Pp.673. ISBN 0-205-07739-0. (Allyn & Bacon: 1981.) Np.
- SHAW READE, S.N. *et al.* (eds). *The Scientific Results of The Oman Flora and Fauna Survey 1977 (Dhofar)*. Journal of Oman Studies Special Report, No.2. Pp.400. ISSN 0379-0703. (Government Adviser for Conservation of the Environment Diwan of H.M. for Protocol, Sultanate of Oman: 1980.) £20.
- SIDDIQUI, M.A.Q., KRAUSKOPF, M. and WEISSBACH, H. (eds). *Molecular Approaches to Gene Expression and Protein Structure*. Proceedings of a Symposium held at the Universidad Austral de Chile, Valdivia, Chile, December 1979. Pp.371. ISBN 0-12-641820-9. (Academic: 1981.) \$27.
- SIMON, N. *Otters*. Pp.46. ISBN 0-460-06974-8. (Dent, London: 1981.) £3.25.
- SIMON, N. *Whales*. Pp.46. ISBN 0-460-06957-8. (Dent, London: 1981.) £3.25.
- SIMON, N. *Wolves*. Pp.46. ISBN 0-460-06975-6. (Dent, London: 1981.) £3.25.
- SMALLMAN, C. *Biology for You, Book 1*. Pp.144. Flexi ISBN 0-09-140891-1. (Hutchinson Educational: 1981.) £2.50.
- SMITH, A.L. *Principles of Microbiology*. 9th Edn. Pp.723. ISBN 0-8016-4682-0. (C.V. Mosby: 1981.) £14.
- SMITH, L.L. *Cholesterol Autoxidation*. Pp.674. ISBN 0-306-40759-0. (Plenum: 1981.) \$75.
- STEINBERG, C.M., LEFKOVITS, I. and DI LORENZO, C. (eds). *Festschrift in Honor of Niels Kaj Jerne on the Occasion of his 70th Birthday*. The Immune System 1, Past and Future. Pp.438. ISBN 3-8055-3407-8. (Karger, Basel: 1981.) SwFr. 248, DM 297, \$148.50.
- SURAN, A., GERY, I. and NUSENBLATT, R. (eds). *Immunology of the Eye*. Workshop III: Infection, Inflammation and Allergy. Pp.300. Flexi ISBN 0-917-000-08-0. (Information Retrieval, Arlington, VA: 1981.) £10, \$25.
- SUTCLIFFE, J.F. and BAKER, D.A. *Plants and Mineral Salts*. 2nd Edn. The Institute of Biology's Studies in Biology, No.48. Pp.66. Pbk ISBN 0-7131-2831-3. (Edward Arnold, London: 1981.) Pbk £2.50.
- TAYLOR, T.N. *Paleobotany: An Introduction to Fossil Plant Biology*. Pp.589. ISBN 0-07-062954-4. (McGraw-Hill: 1981.) £20.95.
- TIPSON, R.S. and HORTON, D. (eds). *Advances in Carbohydrate Chemistry and Biochemistry*, Vol.39. Pp.502. ISBN 0-12-007239-4. (Academic: 1981.) \$55.
- TURNER, M.R. (ed). *Preventive Nutrition and Society*. Proceedings of the British Nutrition Foundation 2nd Annual Conference held at The Royal Society, July 1980. Pp.228. ISBN 0-12-704450-7. (Academic: 1981.) £16.40, \$39.50.
- USSING, H.H. *et al.* (eds). *Water Transport Across Epithelia: Barriers, Gradients and Mechanisms*. Proceedings of the Alfred Benzon Symposium, 15, held at Royal Danish Academy of Sciences and Letters, Copenhagen, June 1980. Pp.501. ISBN 87-16-08600-7. (Munksgaard, Copenhagen: 1981.) D.Kr. 275.
- VON WETTSTEIN, D. *et al.* (eds). *Molecular Genetics in Yeast*. Proceedings of the Alfred Benzon Symposium, 16, held at Royal Danish Academy of Sciences and Letters, Copenhagen, June 1980. Pp.443. ISBN 87-16-08601-5. (Munksgaard, Copenhagen: 1981.) D.Kr. 250.
- WEINER, J.S. and LOURIE, J.A. *Practical Human Biology*. Pp.440. ISBN 0-12-741960-8. (Academic: 1981.) £16, \$38.50.

ANNOUNCEMENTS

Awards

The Rank Prize Funds have announced the following awards: for Opto-Electronics, one prize worth £30,000, goes to **Professor Calvin F. Quate** (Stanford University, California) for his contribution to medical, biological and physical research through the concept of the scanning acoustic microscope, which uses sound rather than light to form images; the second Opto-Electronics prize of £10,000 goes to **Dr Charles Elliott** (Royal Signals and Radar Establishment Malvern, UK) for devising a novel infra-red detector. The Rank Prize for Nutrition is £20,000 and goes to **Dr Hamish Munro** (US Department of Agriculture) for his research establishing the balance of nutrients required as the body grows older.

Arthur C. Clarke (University of Moratuwa, Sri Lanka) has been chosen to receive the eighth Marconi International Fellowship Award in recognition of his scientific achievement for the benefit of humanity in the field of communications science or technology.

The Dr H.P. Heineken Prize for 1982 has been awarded to **Professor C. Weissmann** (University of Zürich) for achievement in the field of biochemistry or bio-physics.

Dr Martin D. Kamen (University of California, San Diego) is the recipient of the 1982 ASBC-Merck Award which recognizes excellence in biochemical research.

Dr Efraim Racker (Cornell University) is the recipient of the 1982 Herbert A. Sober Award for contributions to techniques in biochemical research.

Dr C.W. Daeschner Jr (University of Texas at Galveston) has been presented with the 1981 Sidney S. Kaliski Award of Merit by the Texas Pediatric Society.

The Zoological Society of London has made the following awards for contributions to zoology in 1981: The Scientific Medal to **Professor M.P. Hassell** (Imperial College of Science & Technology) and **Dr J.R. Krebs** (Edward Grey Institute of Field Ornithology, University of Oxford); The Frink Medal for British Zoologists to **Sir Eric Smith**; The Thomas Henry Huxley Award to **Dr N.R. Franks** (University of Leeds); The Stamford Raffles Award to **Lt. -Col. A.M. Emmet**; The Prince Philip Prize to **Jonathan Edward Greenland** (The Grammar School, Bristol).

The biennial Artois-Baillet Latour Health Prize of four million Belgian francs will be awarded in 1983 for an outstanding contribution to the knowledge, the aetio-pathogenesis, the diagnosis and the treat-

ment of cancerous diseases of blood and the lymphatic system both of child and adult. Nominations to: The Secretary General, National Fund for Scientific Research rue d'Egmont 5, B-1050 Brussels before July 1982.

The International Association of Gerontology has created the Sandoz Prize for Gerontological Research. The prize, worth 20,000 Swiss Francs, will be sponsored by Sandoz Ltd, Basel/Switzerland to encourage research in all areas of gerontology and geriatric medicine including biological, medical, psychological, social and other relevant aspects with special emphasis on multidisciplinary research programs. Applications in English before 30 September 1982 to Prof. M. Bergener, Secretary General, International Association of Gerontology, Rheinische Landes-klinik, Wilhelm-Griesinger-Str. 23, D-5000 Cologne 91, FRG.

Appointments

Sir Richard Bayliss is joining the Medical Services Study Group of the Royal College of Physicians of London as an assistant director.

Robert J. Beals (Hall China Co.) is to be the 84th president of the American Ceramic Society.

Sir Ralph Verney, has become a trustee of the School of Water Sciences, High Wycombe, UK.

Meetings

3-6 April, **Medical Consequences of War to the Countries of Europe**, Cambridge (Beer Davies, 1 Melville Rd, Edgbaston, Birmingham, UK).

5-6 April, **Atherosclerosis: Mechanisms and Approaches to Therapy**, London (Dr N.E. Miller, Dept of Chemical Pathology, St Thomas's Hospital Medical School, London SE1, UK).

5-7 April, **Control Systems Concepts and Approaches in Clinical Medicine**, Sussex (The Institute of Measurement and Control, 20 Peel St, London W8, UK).

5-7 April, **The Jubilee Chemical Engineering Symposium**, London (J. Ellis, IChemE, George E. Davis Building, 165-171 Railway Terrace, Rugby, UK).

5-8 April, **Royal Society of Health Annual Congress**, London (Royal Society of Health, 13 Grosvenor Place, London SW1, UK).

6-7 April, **Results and Potential of TMR as Diagnostic and Research Method**, Oxford (Dr P. Hanley, Oxford Research Systems, Ferry Hinskey Rd, Oxford, UK).

7 April, **Analysis of Surfaces and Interfaces in Electronic Devices**, London (The Institute of Physics, 47 Belgrave Square, London SW1, UK).

14-15 April, **Risk, Cost and Pollution**, Oxford (Mr L. Evans, Harwell Education and Training Centre, Harwell Laboratory, Oxon, UK).

14-16 April, **The Assessment of Major Hazards**, Manchester (Mr D.J. Hewitt, 121 Princess St, Manchester, UK).

19-21 April, **3rd International Recycling Congress**, Berlin (IRC, Prof. Dr K.J. Thomé-Kozmiensky, Rhumeweg 14, D-1000 Berlin 37, FRG).

19-21 April, **Cell Behaviour**, London (Prof. R. Bellairs, Dept of Anatomy and Embryology, University College London, Gower St, London WC1, UK).

20-22 April, **Life Sciences — Market Openings**, Sheffield (A. Fraser, Frost and Sullivan Ltd, 104-112 Marlebone Lane, London W1, UK).

21 April, **Biocides in the Oil Industry**, London (Miss I.A. McCann, Institute of Petroleum, 61 New Cavendish St, London W1, UK).

27-29 April, **Esarda Specialist Meeting**, The Netherlands (Mr R.J.S. Harry, ECN, Postbus 1, NL-1755 ZG Petten, The Netherlands).

19-30 April, **Commission for Climatology and Applications of Meteorology**, Washington DC (Research and Applications Dept, World Meteorological Organization, Geneva, Switzerland).

30 April-3 May, **Upper Jurassic and Cretaceous of Normandy**, Le Havre (Dr J. Hancock, Dept of Geology, King's College, London WC2, UK).

30 April, **Electron Beam Damage**, London (The Institute of Physics, 47 Belgrave Square, London SW1, UK).

2-5 May, **Ceramics: A Space Age Technology**, Cincinnati (The American Ceramic Society, Inc. 65 Ceramic Drive, Columbus, Ohio 43214, USA).

4-7 May, **Secretory Immune System**, New York City (The New York Academy of Sciences, 2 East 63rd Street, New York, New York, 10021, USA).

4-7 May, **Council of Biology Editors**, Louisville (P.L. Altman, Council of Biology Editors, Inc., 9650 Rockvill Pike, Bethesda, Maryland 20814, USA).

11 May, **Transient Annealing of Semiconductors**, London (The Institute of Physics, 47 Belgrave Square, London SW1, UK).

11-13 May, **ESTEC Spacecraft Electromagnetic Compatibility Seminar**, Noordwijk (Mr H. Bachmann, Postbus 299, 2200 AG Noordwijk, The Netherlands).

10-14 May, **Oil Pollution Control Course**, Ipswich (Miss I. McCann, Institute of Petroleum, 61 New Cavendish St, London W1, UK).

11-14 May, **Industrial Corrosion**, Geneva (The Center for Professional Advancement, Postbus 19865, 1000 GW Amsterdam, The Netherlands).